For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex libris
universitatis
albertaensis

QUAECUMQUE VERA

Digitized by the Internet Archive
in 2020 with funding from
University of Alberta Libraries

https://archive.org/details/Lindsay1970

THE UNIVERSITY OF ALBERTA

AN APPLICATION OF TWO STOCHASTIC MODELS TO POPULATION MIGRATION IN THE PEACE DISTRICT OF ALBERTA

by

(C) IAN JAMES LINDSAY

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

DEPARTMENT OF GEOGRAPHY

EDMONTON, ALBERTA

FALL, 1970

THE UNIVERSITY OF ALBERTA

AN APPLICATION OF TWO STOCHASTIC MODELS TO POPULATION MIGRATION IN THE PEACE DISTRICT OF ALBERTA

by

© IAN JAMES LINDSAY

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

DEPARTMENT OF GEOGRAPHY

EDMONTON, ALBERTA

FALL, 1970

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "AN APPLICATION OF TWO STOCHASTIC MODELS TO POPULATION MIGRATION IN THE PEACE DISTRICT OF ALBERTA" submitted by IAN JAMES LINDSAY in partial fulfilment of the requirements for the degree of Master of Arts.

Date ....Sept. 15, 1970....

ABSTRACT

Two stochastic models of population migration in the Peace District of Alberta, from 1921 to 1961, have been constructed. One model, utilizing the Monte Carlo method, is basically an experimental sampling device. The other is derived from the Markovian model of migration suggested by Rogers (1966a). The primary aim of the thesis is to examine the relative merits of both models in predicting population distributions.

The theoretical basis of the models assumes an inverse relationship between distance and intensity of local migration. This relationship has been noted in many studies (Hägerstrand 1957, Marble and Nystuen 1963, Olsson 1965a). In the Monte Carlo model, the inverse-distance relationship is incorporated via Pareto expressions. The transition matrices of migration probabilities required by the Markovian model are constructed from gravity models.

Detailed population characteristics are needed by both models. The necessary information was derived from the Albertan "Vital Statistics" reports, the "Census of Canada", and the "Census of Prairie Provinces". A sample questionnaire obtained the migration histories of over six hundred persons, and was used to calculate expected rates of external and internal migration from the twenty-six regions of the study area.

Both models generate a theoretical or expected population distribution for the study area in four different time periods. Observed and theoretical cumulative frequencies are compared by the one-sample

Kolmogornov-Smirnov test. The hypothesis, that any difference between observed and expected values is due purely to chance and is not significant, is either accepted or rejected depending on the maximum difference between the cumulative frequencies and the confidence level chosen.

The Kolmogorov-Smirnov tests indicate that both Monte Carlo and Markovian models reproduce the features of the real world, within reasonable bounds of confidence. The Markovian model is considered more effective than the Monte Carlo method, since natural increase, external and internal migration can be incorporated in a more realistic manner. The calculations required by the Markovian technique are much less laborious than those needed by the Monte Carlo method.

The author adds the qualification that the inability of the Markovian model to incorporate additional factors, other than the independent distance variable, must be overcome. In contrast, the Monte Carlo model modifies the inverse-distance relationship of migration by barriers to movement. However, the performance of the Markovian model, indicates that it could be applied to other forms of spatial interaction, which traditionally have been treated by the Monte Carlo method.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor, Dr. Barr, for his constant advice and understanding. The able assistance of Professor Proudfoot, and the helpful comments of Professor A. Buse, D. Johnson, O.G. Sitwell, and P.J. Smith were extremely encouraging.

Miss Diane Rose prepared the necessary computer programmes, and were it not for her patience and skill in deciphering my instructions, this thesis would not have been written.

I gratefully acknowledge my typist, Mrs. Myron Spak, who managed to interpret my jumbled scrawl, and thank Mr. Jack Chesterman, who reproduced the maps and diagrammes.

Finally, I thank all who helped in any way, not least the residents of the Peace River district of Alberta, who very kindly answered my questionnaire.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

CHAPTER ONE

THE PROBLEM

Aim

The object of this study is to investigate the relative merits of two stochastic models in generating a migration process. The Monte Carlo method is well established as a technique dealing with the diffusion of innovations and the movement of people. Recently, proposals have been made to apply Markov chain models to some of the processes traditionally dealt with by the Monte Carlo method (Olsson and Gale 1968). It is hoped in this study that the relative efficiency of the two models can be gauged by their application to migration in the Peace River District of Alberta from 1921 to 1961.

Each run of the Monte Carlo model produces a simulated population distribution. Since the results of the runs can vary, Pitts (1963) has suggested that the results of a series of simulation runs be averaged. A secondary aim of this study is to evaluate the significance of an averaged simulation solution. There is no similar problem with the analytic Markovian model, which predicts only one population distribution. Brown (1968b) has attempted to synthesize the development of models which treat spatial patterns resulting from spatial behaviour. These applications include models of diffusion of innovations (Hägerstrand 1967a), migration (Morrill 1965b), urban growth (Garrison 1962), settlement (Bylund 1960) and economic development (Curry 1967). The potential of such models has frequently been stressed but the number of studies utilising the Monte Carlo and Markov chain

techniques is still relatively small. The performance of the Markov chain model in the limited field of migration has important consequences.

The general nature of these models is emphasized. Each model contains facets which are worthy of further study. The correct use of distance-decay functions in Monte Carlo methods has aroused considerable controversy (Morrill 1963a). Testing of simulated results against real world observations raises the question of which non-parametric tests to use (Pitts 1963). While these problems have been encountered and dealt with, the accent in this thesis is on the achievement of a general working model rather than on suggestions of detailed modifications to the operation of the models. The validity of this approach is considered reasonable, since the application of such models has lagged behind suggestions about their utility.

The process of migration in this study has been considered in specific terms. Explanation of population distribution and redistribution should take cognizance of a variety of physical, economic, social and political factors. Only migration—the way in which redistribution of population occurs—has been treated here. Conditions of natural increase among the settled population and external movements are constraints given to the models from outside (Morrill 1965b, p. 101). Local migration is treated internally and permitted to vary according to the principles of the models. The intensity of local migration is related to information, the supply of which falls off with distance (Morrill and Pitts 1967, Marble and Mystuen 1963). This is low-level explanation. If detailed explanations of spatial patterns are sought, then space preferences have to be examined. Spatial patterns are the result of individual or collective

comparison and evaluation of alternative spatial opportunities (Rushton 1969). There is no reason to suppose that higher-level explanation, involving human perception, can not be built into the basic framework examined here.

Theoretical Background

1. Models

The author is concerned with processes which create spatial patterns. The place of the model in the Scientific Method is proven, though the term model has many connotations. Chorley and Haggett (1967) have set out in great detail the characteristics and functions of various types of models. The model is commonly regarded as a simplified picture of reality in which certain relationships are generalized. In this sense the model is an approximation, and its value must always be set in regard to its degree of abstraction. Such models are educational, because phenomena can be visualized and comprehended in a fashion not otherwise possible. Insight into the process of migration is gained explicitly in this study through the operation of the models, and implicitly through the labour of fashioning them. Models are also a means of organizing data. They provide frameworks into which information can be fitted and ordered. Models act logically and account for the behaviour of phenomena, since explanation involves the collapsing of complex systems into simpler ones where interdependencies and relationships can be observed.

To many gecgraphers, the model is a means to link theory and observation. Models are abstractions, and the value of the model is

conditioned by the degree to which it approaches and corresponds to the real world. Depending on the researcher's aim, models can be used to make predictions about the real world that they model. In this sense, the model tests a hypothesis or expectation based on theoretical postulates. If the model and the real world agree within the bounds of accepted limits, then the hypothesis or the postulates can be accepted as valid or at least not rejected. Even if there is considerable difference, extensions and modifications to the model should suggest themselves.

The models in this study are based on the assumption that there is a relationship between migration intensity and distance. If the generated results compare "closely" with the real world then this inverse-distance relationship can be considered "true". But how should the modelled results be tested against the real world? When does any difference become significant? Such questions are considered more fully in a later chapter. Many studies involve reasoning about the theoretical postulates, while neglecting the important aspect of testing these assumptions (Pitts 1963, pp. 116-119).

The simple inverse-distance postulate has been modified by space preferences, such as the operation of physical or functional barriers to movement. The models in this study are probabilistic rather than deterministic, because they relate to situations where not all of the variables or relationships are known. The models treated here are historical-predictive, since they deal with a process operating through time and attempt explanation of that process.

2. A Stochastic Viewpoint

Until recently most work in Theoretical Geography has been deterministic as a result of the importance in geography of the optimising notions of Economics (Olsson and Gale 1968, p. 229). Typical of such concepts was the hypothetical "Economic Man", who had perfect knowledge and the capacity to make completely rational decisions. Curry (1966 p. 42) described the deterministic approach as the "specification of variables and causes of certain intensity and interaction which will lead to a result differing from reality by an error term". If the problem of predicting a pebble's movement downstream were considered, or the difficulty of specifying a migrant's distination, then the immensity of this task become evident. All variables and causes can not be specified. A deterministic model of pebble or migrant behaviour could be constructed but, given man's level of ignorance and the multiplicity of factors involved, such models would lack reliable predictive power. If a deterministic approach were adhered to, then the problems that geographers can deal with would be effectively limited. Attempts to improve the predictive power of their models have led geographers to conceive behaviour as probabilistic or random and to construct[1] probabilistic models.

Examples of the probabilistic approach can be found in the Physical Sciences (Cole and King 1968, pp. 98-99). The deterministic relationship of cause and effect has been rejected in Quantum Theory. Elemental particles are considered to obey laws of chance, and their

---

[1]Random behaviour has the implication of complete haphazardness. Probabilistic or stochastic are preferred terms.

behaviour can be expressed only in terms of statistical probability. On the macro-scale of the universe, the Theory of Relativity also discredits determinism. Since position and time are not constant, effect possibly could precede cause. At earth scale, where the law of gravity holds, the precedent of Quantum Mechanics and Relativity becomes less evident. Yet, by following probabilistic reasoning, more is achieved by accepting less.

For both pragmatic and philosophical reasons, geographers have been encouraged to attempt probabilistic explanations. A process is regarded as stochastic if it "does not proceed according to any immutable law but is at least partly dependent on random and chance factors" (Bartlett 1962, p. 49). At this stage, for the purpose of clarity, a distinction should be made between "grand probability"—the consideration of degrees of uncertainty in treating behaviour—and Probability Theory—the essential mathematical basis of statistical theory[2]. Stochastic models are developed by reference to Probability Theory.

There are many fields in Geography where the researcher can appreciate a viewpoint in which decisions and actions are set within a probability framework. Some studies have treated individual behaviour in probabilistic terms (Wolpert 1965), and others have dealt with group responses (Golledge and Brown 1967). Attempts have been made to phrase traditional concepts in new terms, such as Curry (1967) who suggested a new theory of central place systems in terms of queuing theory.

---

[2]Some people regard the Theory of Probability as a branch of pure mathematics, and Statistical Theory as the application of this mathematical theory to statistical phenomena. (Bartlett 1962, p. 15).

Probability models have been developed on an aggregative level for intra-urban (Malm, Olsson, and Wärneryd 1966) and inter-urban systems (Morrill 1963b).

Whether the problem is physical or human, spatial or temporal, individual or collective, retrospective or futuristic, the probability approach has validity (Curry 1966).

3. Simulation Versus Analytic Approaches

A probabilistic approach can solve some but not all problems in geographic research. Among successful applications of Probability Theory in Geography is Dacey's work on the analysis of point patterns (1966). Yet, such efforts are essentially descriptive rather than explanatory. The imprecise nature and great number of variables in geographic problems has frequently resulted in models that are over-simplified. Analysis through the use of means and dispersions, regression, and correlation approaches have been employed in many studies. Unfortunately, standard statistical techniques need not provide required answers. When over-simplified models are used, little more than a crude comparison with actual data may be justified.

The difficulty of deriving adequate analytic solutions has encouraged many researchers to accept methods of simulation in place of more formal mathematical theory. Of the former, the Monte Carlo method is a widely used technique. Its characteristics will be examined in Chapter Two. Simulation approaches, providing methods of analysis where no others are possible, are not always satisfactory because of the difficulty of evaluating the results. In contrast, the Markov chain

model presented in Chapter Three is an analytic solution.

Method

To consider temporal as well as spatial characteristics, the models were applied to four ten-year periods from 1921 to 1961. The procedure consisted of:

1. collection of appropriate information and its transformation;
2. creation of Monte Carlo and Markov chain models;
3. operation of the models;
4. analysis and evaluation of the results.

1. The Derivation of Information

The Peace River study area was divided into twenty-six regions, each of which was an amalgam of census sub-divisions. For each region, in each time period, information was collected on population numbers and density, natural increase, number of people available for internal migration, migration characteristics, and distances between all regions. In addition, certain barriers to movement between regions were determined.

2. Creation of the Models

The derivation of the models is treated in detail in Chapters Two and Three. At this point, it is worthwhile to consider some peculiar difficulties. The redistribution of population is only one aspect of a general stochastic model of a migration process. Assume a certain population in a region at a time t. In the following period, whatever its duration, some people will leave, some will die, others will arrive or be born. The composition and size of the region's

population will change. The elements of the change, births, deaths, and migration, occur together continuously. Any operational procedure must first segregate change into discrete time periods. Secondly, the elements of change must be treated separately. Conceivably, a model could be designed to avoid these simplications, but its complexity and the amount of data required would be prohibitive. The only feasible alternative is to design operational modes which are nearer to, than further from, the truth. Consequently, the Monte Carlo model segregates natural increase, external and internal movements. In this respect, the Markov chain model is more flexible, although it too requires the different components to be introduced separately. This point will be taken up again during consideration of the models' design.

3. Operation of the Models

The objective of both models was to create a new population distribution at the end of a time period, given the distribution at the beginning of the period and certain operating procedures. Natural increase and external movements are taken as constants, but local migration varies according to the principle of the model. The latter depends upon the probability of interaction between regions, which is derived theoretically. Since the number of people moving in any one time period was great, and the probability of interaction extremely complex, the models are designed to be run by computer[3].

4. Analysis and Evaluation of Results

The testing of generated results against actual observations is

---

[3]There are a number of programmes relevant to these studies including Marble (1964) and Pitts (1965 and 1967). The programmes necessary for this study were prepared by Dianne Rose.

a crucial aspect of all models. Unless this procedure is adequate, the model will lose any claim that it has to objectivity. The assumption is made that the real world has only a probability existence. It is only one of a possible series of results. Consequently, there is no information available on the nature of the underlying distribution from which it comes. Since no assumptions, such as normality, can be made about this underlying distribution, non-parametric tests of correspondence have to be used.

The results of the models have been tested, for each time period, against each other, and against the real-world situation.

The Study Area

The study area is the core of the Peace River District of Alberta (Figure 1). Settlement in its northern and southern districts have been excluded because of the small size of their population. The operation of the models was observed through time. The Peace is a laboratory in which the process of frontier settlement and population redistribution can be examined almost from the beginning of settlement. The area may be compared to a newly emerged volcanic island, which goes through the complete stages of growth from pioneers to associations. The study area is described in greater detail in Chapter Four.

Data Base

Data sources do not compare well with Sweden where many Monte Carlo models have been applied (Hägerstrand 1957, Morrill 1965b).

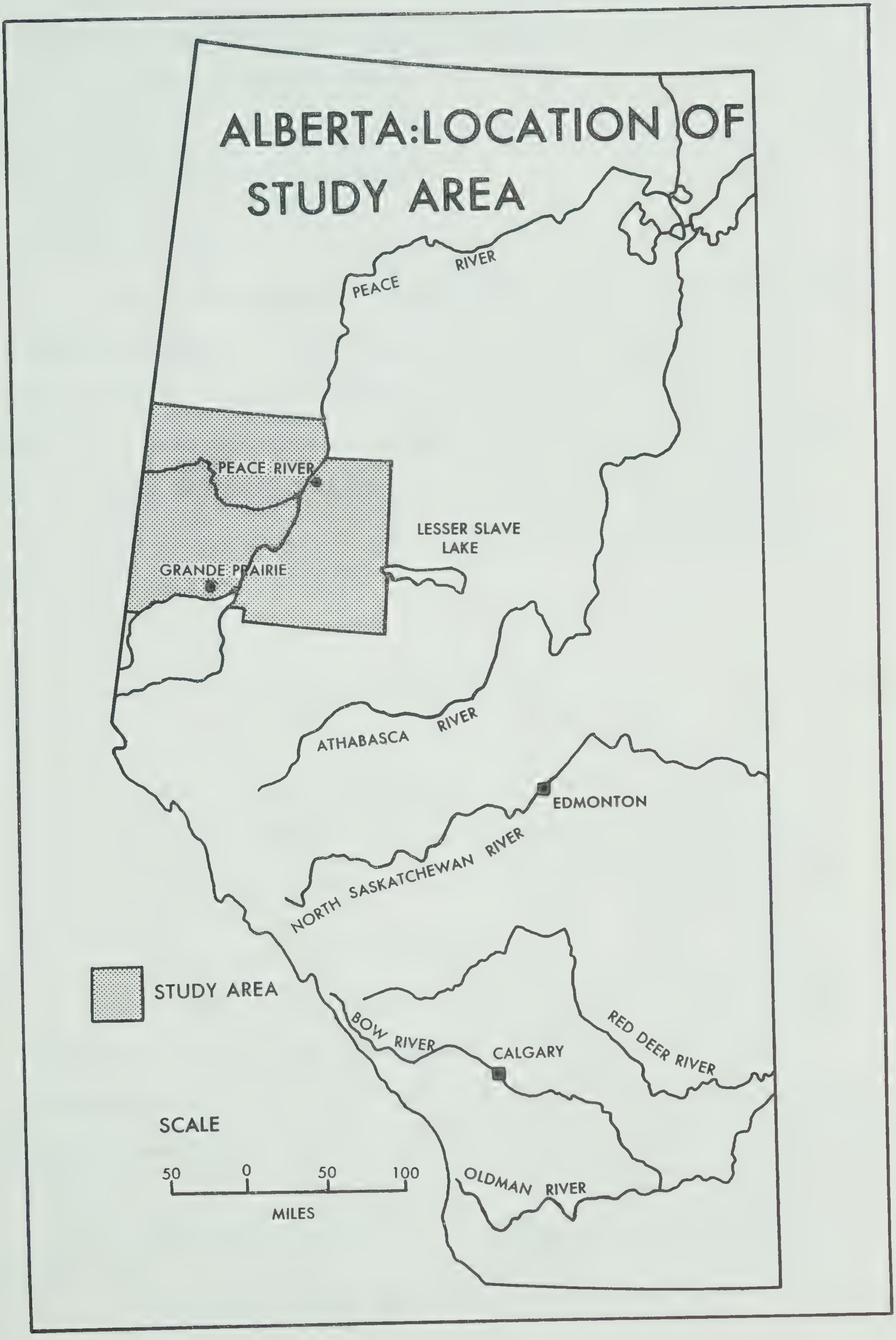

ALBERTA: LOCATION OF STUDY AREA
PEACE RIVER
PEACE RIVER
LESSER SLAVE LAKE
GRANDE PRAIRIE
ATHABASCA RIVER
EDMONTON
NORTH SASKATCHEWAN RIVER
STUDY AREA
BOW RIVER
CALGARY
RED DEER RIVER
SCALE
50 0 50 100
MILES
OLDMAN RIVER

FIGURE 1

Complete disaggregated data on population and other characteristics are absent. Better data would result if a shorter, more recent time period was taken. Yet, investigation of the performance of the models through time is considered important.

1. Written Sources

The written sources of information are the Dominion Bureau of Statistics censuses for Canada and the Prairie Provinces[4], and the Albertan "Vital Statistics" Reports. The study area is composed of census divisions fifteen and sixteen[5]. Births and deaths are recorded in the census for divisions. Consequently, the net migration gain or loss for the entire study area is obtained by the residual method (Isard 1960, pp. 54-55),

$$M_\theta = (P_{t+\theta} - P_t) - N_\theta \qquad (1.1)$$

where M_θ is the net migration over the time period, P_t the population at the beginning of the time period, $P_{t+\theta}$ the population at the end of the time period, and N_θ the natural increase over the time period.

The same procedure was used for the regions over the period 1932 to 1944, since births and deaths are given at the census subdivision level in the Albertan "Vital Statistics" for these years (but not before or after). For the rest of the time, formula (1.2) was used,

---

[4]The Separate Prairie Provinces Census was discontinued after 1946.

[5]After 1951 the two divisions were combined.

$$M_\theta = (P_{t+\theta} - P_t) - (\alpha P_\theta + \beta P_\theta) \tag{1.2}$$

where M_θ is the net migration over the period, P_t is the population at the beginning of the period, $P_{t+\theta}$ is the population at the end of the period, α is the birth rate over the period, β is the death rate over the period, and P_θ is the population of the region to which birth and death rates are applied. Crude birth and death rates were used since no breakdown by age and sex is given for census sub-division.

Resultant figures for regions are reasonable approximations of the true situation. Despite their limitations, the data are adequate for the general models in this study.

2. The Sample Questionnaire

The number of potential migrants in each region was derived from information gained from sampling with a questionnaire. Many of the original settlers or their first generation descendants still survive. This was an important consideration in deciding to sample the population.

The sample population was based on the Peace River Telephone Directory which listed 12,235 private households. One hundred and fifty were drawn completely at random, with a supplementary list of fifty. A sample size of one hundred and fifty was chosen over the conventional figure of one hundred, since, at one stage the transition matrices of migration for the twenty-six regions were to be estimated from the sample. To ensure complete coverage for all the six hundred and seventy-six cells, the larger sample size was taken. In practice,

the transition matrices were estimated by a gravity model, not empirically from sample information.

Interviews on the migration history of individuals in the households were conducted by telephone and mail. In this way a 98% response was achieved. Each household was sent a copy of the questions to be asked (see Appendix D), and informed that a telephone call would follow. Some questionnaires were completed and returned by mail, but the bulk of the interviews were carried out over the phone. The high degree of success was obtained by use of the supplementary list when individuals had moved, were on vacation, or no response was achieved after several telephone calls. Since migration was being considered back to 1921, individuals were asked to provide information, not only about themselves and their spouses, but also their parents and in-laws. A complete spectrum of knowledge of the movements of six hundred and fourteen persons over the period 1921 to 1961 was achieved.

The sample did not include interviews with people who lived in the Peace prior to the date of interview and who left for other regions of the world. Similarly, people who may have lived in the Peace for several decades, for example, and then died without leaving offspring in the Peace River, could not be included. Furthermore, some bias may have been introduced into the sample because interviews were arranged only with those families which were listed in the telephone directory. The sampling procedure might introduce a bias, if people with telephones were more mobile than those without. On the other hand, there is a possibility that people with telephones might be better adjusted to the

local area, and less likely to migrate. These forms of bias could not be more adequately determined in this study.

Finally, the size of the Peace Region might vary from that of areas used by other researchers. It is important that the results of this study be interpreted within the general size and situation of the Peace Region. The size of the sub-regions was determined by the desire to reduce the total number of census subdivisions to a manageable number. The aggregation of sub-units into twenty-six larger districts was determined on the basis of ethnic, topographic, and settlement characteristics of the census sub-divisions of the Peace District.

In conclusion, the primary aim of the study is to investigate the relative advantages of the Monte Carlo method and the Markov chain model. The results achieved by the models for each of the four decades over the period 1921 to 1961 are tested against the actual population distribution in the Peace River study area.

The history, nature, and detailed construction of the models are given in Chapters Two and Three. Chapter Four deals with the operation of the models, while predicted values and observations are tested against each other in Chapter Five. The results of the tests are analysed and evaluated in the final chapter.

CHAPTER TWO

THE MONTE CARLO METHOD

This section describes the history, potential, and construction of simulation models. The terms, simulation and model, could be regarded as synonymous, though in practice, the former refers to a specific group of models, of which Monte Carlo is the best known. Monte Carlo simulation treats situations where not all the factors and relationships are known. Hence, a probability concept is introduced. Basically, Monte Carlo methods are concerned with sampling experiments on random numbers[1].

History of Monte Carlo Methods .

The term Monte Carlo, which formally emerged in 1944, has various aspects. Its significance to a mathematician is different to that of a physical scientist or a geographer. Partly, this is due to the method's threefold origin (Tocher 1963, p. 2).

The oldest application of Monte Carlo methods lies in the sampling experiments of statisticians. While there were numerous sampling experiments in the second half of the 19th century[2], modern

---

[1]Consider, for example, the problem of estimating the area of an irregular region. One method would employ a sampling experiment. The irregular region is surrounded by a region of known area. One hundred points are located by a random procedure over the known region. The proportion of points which fall in the irregular region provide an estimate of its size as a proportion of the known region's area.

[2]There were numerous experiments of throwing a needle in a haphazard fashion on to a board ruled with parallel straight lines. The value of π was infered from the number of intersections between needle and lines.

sampling experiments began with the work of "Student" (W.E. Gossett). In 1908 he used sampling to bolster his faith in the "t distribution" which he had derived by theoretical analysis. This work was taken up by Pearson, Galton, and others (Tocher 1963, pp. 43-44).

The object of sampling experiments was to determine the amount and form of variability that some estimate of a parameter would have in repeated sampling. Although, mathematical techniques have developed to great levels of sophistication, the value of sampling experiments in mathematical statistics still remains, especially with the availability of high speed computers to perform laborious calculations.

The second origin of the method was created by the demands of applied mathematicians and physicists. In connection with work on the atomic bomb during World War II, von Neumann and Ulam utilized the technique to simulate the probabilistic nature of neutron diffusion in fissionable material. Since then, the technique has been employed to deal with many physical systems where particles behave in a partly regular, partly irregular fashion (Meyer 1956). The Monte Carlo method permits a participant to play a game of chance, the rules of the game allowing deterministic and stochastic features of the process to be reproduced. By considering the results of a large number of trials, the distribution of particles or the number of particles which escaped through a medium at the end of a time period can be estimated. Similar techniques have been used to determine when particles are going to collide, the number of resultant particles and subsequent collisions. Monte Carlo, in this sense, is a "device for studying an artificial stochastic model of a physical or mathematical process", and involves the

use of sampling within a probabilistic framework (Meyer 1956, p. 5).

Physicists have extended the technique to deal with deterministic problems (Hammersley and Handscomb 1964 , p. 7). Partial differential equations arising in the abstraction of physical processes can not always be solved by conventional analytic methods. It is possible to find a probability analogue to the deterministic equation. The Monte Carlo method enables the statistics of the analogue to be computed by a sampling experiment. An answer is thus provided for the deterministic equation.

The development of Monte Carlo techniques in Operations Research, the third origin, is most relevant to Geography. Operations Research seeks to apply the Scientific Method to military, managerial, and industrial problems. In keeping with many sciences, it sets up models to provide insight into processes considered. Unlike Physical Science, but similar to Geography, Operations Research studies a great variety of phenomena. Situations arise which, when incorporated into models, can not be solved adequately by normal mathematical methods. Many researchers, like applied mathematicians, have turned to sampling experiments to achieve results where no other methods are possible (Harling 1958).

Simulation techniques have passed via Operations Research into the Social and Behavioural Sciences, including Geography. There, too, they offer an approach where no others seem likely.

Owing to the multiple origins and use of Monte Carlo, the

technique can be classified in different ways. The simplest system is a division into "probabilistic" and "deterministic" Monte Carlo (Hammersley and Handscomb 1964, pp. 1-12). The former serves as a device to study an artificial stochastic model of a physical or mathematical process. Random numbers are chosen in such a way that they directly simulate the random events in a stochastic process. The desired solution is inferred from the behaviour of the random numbers. In the simplest and crudest "direct" simulations, no more is involved than transference of the stochastic features of the real world into the model. In more sophisticated efforts only those features of the real system are incorporated which affect the solution of the model. In other words, care is taken to ensure that the model neither oversimplifies the system, nor incorporates so many elements that it becomes unwieldy (Harling 1958, p. 307).

Deterministic Monte Carlo involves the creation of probability analogues to solve deterministic problems. In this context, Monte Carlo is used to fill gaps in theoretical mathematics or deterministic reasoning.

Since most "geographic" processes are regarded as stochastic, Monte Carlo techniques in Geography typically are of the probabilistic type. Bunge (1969) has suggested that simulation and conventional analytic techniques could be combined. A form of "deterministic" Monte Carlo could be employed in which, for example, parameters of probability functions could be estimated. As yet, this is an unexplored avenue of research.

The Monte Carlo Method in Geography

Much of the credit for introducing the method into Geography must be given to Hägerstrand. The impact of his book, published in 1953, has been so considerable, despite a time lag caused by lack of translation into English (Hägerstrand 1967a), that few fields relevant to Geography are now untouched. Monte Carlo methods have been applied to the diffusion of innovations (Hägerstrand 1965), migration processes (Morrill 1965b), growth of urban settlement (Morrill 1960), and intra-urban growth and differentiation (Malm, Olsson, and Wärneryd 1966). The popularity of the technique has led to efforts to develop a formal basis (Brown 1968a). As Pred suggested (1967), applications range from analysis of past distributions, future projections, and treatment of static distributions.

Essentially, these methods have been employed to deal with problems where no satisfactory mathematical solutions are feasible. This is the case when the prediction of a migrant's move is considered. A deterministic approach must be replaced by a probabilistic one, if a greater degree of precision is to be achieved. Admittedly, the transformations of space suggested by Tobler and Bunge (Bunge 1969) may reduce the necessity of so many random variables in model development, though it is unlikely that such transformations could make a random approach completely feasible. Equally, with many problems in Geography, conventional Probability Theory can be applied only to crude and over-simplified models.

The last comment can be appreciated by considering a simple analogy. Suppose that it is desired to know the chances that the

number of hits in ten throws at a coconut shy is even. This could be achieved by taking a series of ten thousand throws at the shy. On the other hand, conventional probability theory can provide an answer without the necessity of experimenting (Schreider 1966, p. 3). However, if it were desired to estimate that the chances of a given game of partience coming out were greater than even, then there is no way that Probability Theory can provide a reliable answer. The number of possibilities for any one trial is too great. The only way to achieve an estimate is to play the game again and again (King 1951, p. 13). This applies to all processes, including migration, where the phenomena can vary greatly.

A crude deterministic model of migration, such as the gravity model, could be derived, or probability functions could be developed. The level of precision is low because of oversimplification. It is here that the value of the experimental Monte Carlo method can be appreciated. Underlying theoretical considerations, such as the inverse-distance relationship, can be incorporated. Yet, the allocation of random numbers to random events permits the element of chance to be introduced as well, so that the researcher's level of ignorance and internal irrational behaviour of the system can be appreciated. A simple example will clarify these points.

If the intensity of flows or interactions between points is assumed to fall off with increasing distance, then this relationship may be expressed in a simple statement, such as the gravity model.

$$I_{ij} = g \frac{P_i \cdot P_j}{D_{ij}^b} \tag{2.1}$$

where I_{ij} is the interaction between place i and place j, P_i and P_j are the populations of i and j respectively, D_{ij} is the distance between i and j, and g and b are empirically derived constants.

The gravity model is a deterministic statement. It can be converted to a probabilistic expression, such as the Pareto function,

$$p(d) = \frac{c}{D^b} \tag{2.2}$$

where $p(d)$ is the probability of interaction between two points, D is the distance between, and c and b are empirically derived constants. Of course, several such expressions could be developed and their relative merits argued (Morrill 1963a).

This is the starting point of the simulation. In some cases, equation (2.2) is fed directly into the model (Pitts 1967), but more often a special probability matrix is employed (Figure 2a). Conventionally, this is a five by five grid, where each cell is 25 square kilometres (Morrill and Pitts 1967). The operational procedure superimposes the grid over each potential migrant in turn, so that the migrant's location becomes cell 13 (Figure 2a), the centre of the grid. The probability of movement to any other cell is given by the figures in each cell (Figure 2b), which sum to unity. These probability figures are derived

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Figure 2a

The Probability Matrix

.022	.028	.032	.028	.022
.028	.045	.063	.045	.028
.032	.063	.127	.063	.032
.028	.045	.063	.045	.028
.022	.028	.032	.028	.022

Figure 2b

Probability of Interaction

000-021	022-049	050-081	082-109	110-131
132-159	160-204	205-267	268-312	313-340
341-372	373-435	436-563	564-626	627-658
659-686	687-731	732-794	795-839	840-867
868-889	890-917	918-949	950-977	978-999

Figure 2c

Random Number Targets

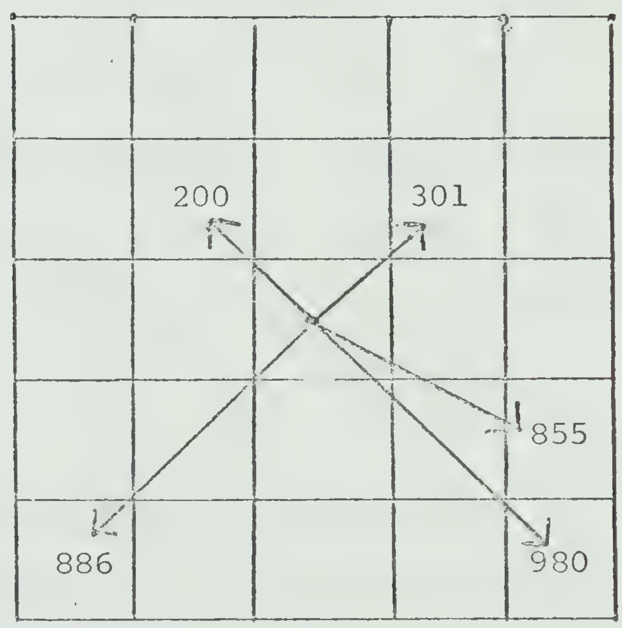


Figure 2d

Destinations of Migrants

FIGURE 2

MONTE CARLO PROBABILITY GRID

Source: Malm, Olsson, and Wärneryd (1966)

from expression (2.2). Since most probability functions tend to give inaccurate estimates for close-in distances (Morrill and Pitts 1967, p. 415), the value of cell 13 is customarily derived from empirical observations. The other probabilities are then adjusted accordingly. It should be apparent that the cell values run symmetrically about the grid, such that the values for cells 18, 14, 8, and 12 are similar, and so on (Figure 2b).

The Pareto function, or some similar expression, illustrate no more than a crude relationship. Each migration can be regarded as a random event and can be allocated a random number. If the probabilities in the matrix are converted to a tally of discrete numbers (Figure 2c), a range of digits is derived for each cell. Since there are more numbers in cell 13 than in any other cell, the possibility of selecting a random number in that range is greatest. Yet, because numbers are drawn completely at random, there is always the possibility of a less likely destination being chosen. Assume that five migrants are to be distributed by the illustrated system (Figure 2c), that the five draw numbers 200, 301, 855, 980, and 886, and hence move to locations 7, 9, 20, 25, and 21 respectively (Figure 2d).

By using similar techniques physicists have estimated the distribution of elemental particles. Similarly, in the migration process, the method permits the distribution of migrants to be estimated at the end of a time period. Each migrant is assumed to behave in a partly regular, partly irregular fashion, the former through the distance-decay function or some similar theoretical postulate, and the latter by the allocation of random numbers.

This is an extremely simple illustration. Other considerations need to be raised, and will be discussed in the next section. In summary, simulation provides three basic advantages to geographic research.

1. Insight into the Process Concerned

The experimental running of a process can often be heuristic. Simulations of innovation diffusion have suggested refinements to the idea of a diffusion wave, such as the filtering down through the urban hierarchy (Hudson 1969b). Such simulations, operating like time-lapse photography, have emphasized the continuing importance of the core area (Hagerstrand 1965, p. 56). The ability of the technique to act as a framework for the organization of data suggests that this should be its main function.

2. Pragmatism

A number of studies in Geography and Urban and Regional Planning emphasize the value of the technique as a tool to envisage the impact of change upon a situation such as urban growth. (Goldberg 1968; Recht 1968; Walter 1968). Morrill (1965c) has used the method to suggest the expansion of the Negro Ghetto. Elsewhere, the technique has been used to illustrate the effect of changing urban renewal inputs upon urban growth[3].

3. Testing of Hypotheses

If probabilities are derived on the basis of certain postulates, then the simulated distribution should reflect the theoretical postulates.

---

[3]For example, note "Community Renewal Programming: A San Francisco Case Study", a report prepared by Arthur Little Inc. in 1966.

If the results are tested against the real world, then the degree of correspondence should indicate how accurate the postulates are. In practice, serious difficulties arise with regard to the testing of simulated patterns, so that the heuristic and pragmatic qualities are often claimed as the technique's chief attributes. Chapter Five treats the difficulty of deriving adequate tests in detail.

Some Relevant Considerations in the Construction of a Monte Carlo Model

The chief virtue of the Monte Carlo method is that it permits ignorance to become a strength. While less vigorous than mathematical approaches, it offers a solution where no others are, as yet, conceivable. There are, however, several operational pitfalls.

1. Number of Observations

Since the method is based on the selection of random events, one simulation can vary greatly from the next. The implications of this for evaluation will be discussed below, but this characteristic also implies that a large number of independent events must be necessary. It seems appropriate only to apply the technique to problems where there are a large number of events. A large number of people move at any one time during migration in the Peace.

2. Selection of Appropriate Relationships

Garrison (1962) suggested that the application of the simulation technique should be carefully chosen, lest the researcher's original ignorance be reproduced. If the unlikely task of simulating all possible relationships in the real world succeeded, then all that would have been produced is a "black-box" model. This would provide no insight into how

the processes function. Instead, it is necessary to concentrate on certain relationships within a simplified framework, and thereby, through continual modification and verification, to build up knowledge. Thus,

> Simulation ... does not ask for exact replication, for that would destroy or obscure the generality of the model and the theory on which it is based. Of course simulation will reflect the rules of the model itself, but owing to randomness, or freedom of choice, a range of variability may be expected. In general, we wish a simulation to be enough like the real world that both could have been produced from the same process (Morrill 1965b, p. 97).

3. Theoretical Significance of Probabilities

Closely associated with the last point is the derivation of probabilities of random events. The real world could conceivably be reproduced exactly in a simulation by carefully modifying the probabilities through a feedback process. Equally, some set of probabilities could be derived by nothing more than a casual examination of the process. Unless theoretical meaning can be given to the probabilities, Monte Carlo methods become crude direct simulations of the process concerned. Little or no insight into the process is gained. Hägerstrand (1965) was able to provide the probability of innovation diffusion or migration movement with theoretical significance, by relating the interaction to a distance-decay function expressed in the Mean Information Field concept (Marble and Nystuen 1963).

4. Dynamic Rather Than Static Situations

These techniques are frequently employed to treat dynamic or changing situations. Such is the case with historical-predictive models. Consequently, probabilities should not be expected to remain constant over time, especially when human rather than physical processes are

treated. Changing inputs or feedback will result in changes in the system. Conventional approaches derive different sets of probabilities for different time periods, though of course, time changes are continuous and not discrete.

There is also the problem of equating a simulation run, or a generation in a simulation, with real time. In diffusion studies the usual method relates simulation to time by means of a linear-logistic curve (Dodd 1955), since the number of adopters in a population follows a double asymptotic or S-shaped curve through time[4]. Intuitively, it seems more difficult to apply this concept to the migration process.

5. The Significance of a Simulation Run

The problem of gauging how well the theoretically simulated pattern corresponds with the real world is closely related to that of judging how well one simulation fits the theoretical distribution. Does one simulation provide a good indication of the distribution based on theoretical postulates, if it is known that simulated results vary widely? If not, then how many simulations are necessary? If an "averaged" solution were derived, would it be valid?

The simulation distribution, a statistic which estimates a population parameter, represents one pass through an unknown, underlying population. As with all sample statistics, the question is asked,

---

[4]The number of adopters is at first small, then increases rapidly, and finally falls off as saturation of the population is reached.

"how representative is the statistic of all observations that might be made?" Good sampling practice involves the use of variance-reducing techniques, the specification of an adequate size of sample, and an estimate which indicates how likely the sample is representative (Hammersley and Handscomb 1964, p. 5). Unfortunately, the nature of the irregular, simulation statistic and lack of knowledge of the underlying population do not permit the use of conventional sampling practice.

Error terms can not be derived. In some problems in Operations Research, variance-reducing techniques are employed, whereby certain transformations and alterations to the original data are carried out to increase the truth of the result (Tocher 1963, p. 116). Application of such techniques has not been possible in the cases where the method has been employed in geographic studies. Although alteration of probabilities by feedback could be envisaged as a form of variance-reduction, the theoretical and empirical significance of the data might be ruined.

In the absence of variance-reducing techniques, a good approximation of the theoretical population is achieved by increasing the size of the sample (Tocher 1963, pp. 2-3). Since the accuracy of estimates for the parameters of the unknown distribution is a function of the number of observations, computer programmes have been written to "average-out" a large number of simulation results (Marble and Bowlby 1968). The suggestion is that the "averaged" solution more closely approximates the ultimate or "omega" distribution than one simulation (Pitts 1963). Yet, how many simulations must be averaged

before a good approximation is derived? Maass et al. (1962), in considering the design of a water resource system, have suggested a close fit between the ultimate distribution and a simulated solution after two hundred runs out of a possible 10^{23} simulations[5].

This is by no means a satisfactory solution, nor is it feasible in all cases. Certainly, if simulation runs are regarded through time, an "averaged" solution may approximate the omega distribution. Much depends on the point of view of the researcher, and the use to which he puts his work. If history is seen in terms of degrees of events rather than particular events (Curry 1966, p. 46), then this solution appears adequate[6].

On the other hand, if the results of a simulation are viewed spatially, then the significance of an "averaged" solution is difficult to envisage. Morrill (1965b, p. 172) suggested that "we may not average unique patterns". A pattern could be simulated over and over again, but if the results are averaged, all that is achieved is a smooth spatial surface. For example, if two points on a transportation system are taken, then the number of simulated links between them can not be averaged out. The Monte Carlo simulation model employed in this study produces a new population distribution at the end of each time period. This distribution, like that of the real world, can be regarded as one sample from many possible solutions. Since discrete population units are treated (rather than network links) an "averaged" solutions seems

---

[5]In this case, the number of variables is very much smaller than in the case of migration in the Peace District.

[6]If degrees of events are accepted, then real history becomes only one outcome in a whole series of possible outcomes.

valid.

6. Inverse-Distance Relationship

The distance-decay relationship has been summed up by Zipf's "Principle of Least Effort" (1949) and Bunge's (1969) contention that things locate as close to each other as possible. An early deterministic summary of the relationship was the gravity model. More recently the concept has been embodied by distance-decay functions, such as the exponential, Pareto, or log-normal.

The inverse-distance relationship provides a basis for many studies including diffusion (Hägerstrand 1967a), migration (Morrill 1965b), urban growth (Malm, Olsson, and Wärneryd 1966), and other forms of interaction such as consumer behaviour (Reilly 1931). For example, in the absence of adequate data on social contacts, Hägerstrand (1957) fell back on surrogates such as local migration and telephone calls. Basically, Hägerstrand assumed that the spread of an idea occurs largely as a result of personal contacts. If the spread is governed by information, then there is clearly a relationship between information and distance. Time, cost, and effort of interaction become greater in absolute terms as distance increases.

Migration falls off with distance (Hägerstrand 1957). Furthermore, migration is a function of information, particularly when the influence of relatives is considered. Recently, the focus has shifted from the distance-decay relationship. The personal evaluation of spatial opportunities has been proposed as a more critical concept (Rushton 1969). However, the importance of distance should not be too

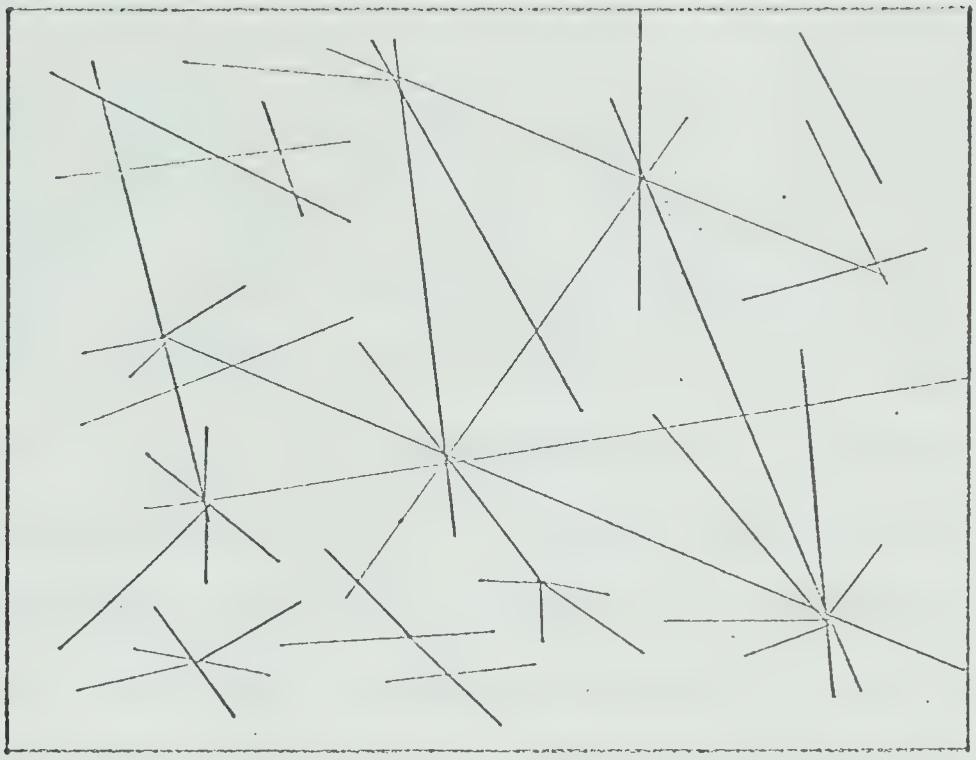
easily discarded. Individual perception of distance would seem to be a crucial aspect of any personal evaluation of spatial opportunities. A more formal appreciation of distance is required, such as the derivation of some functional distance, rather than straight line measurement.

7. The Mean Information Field

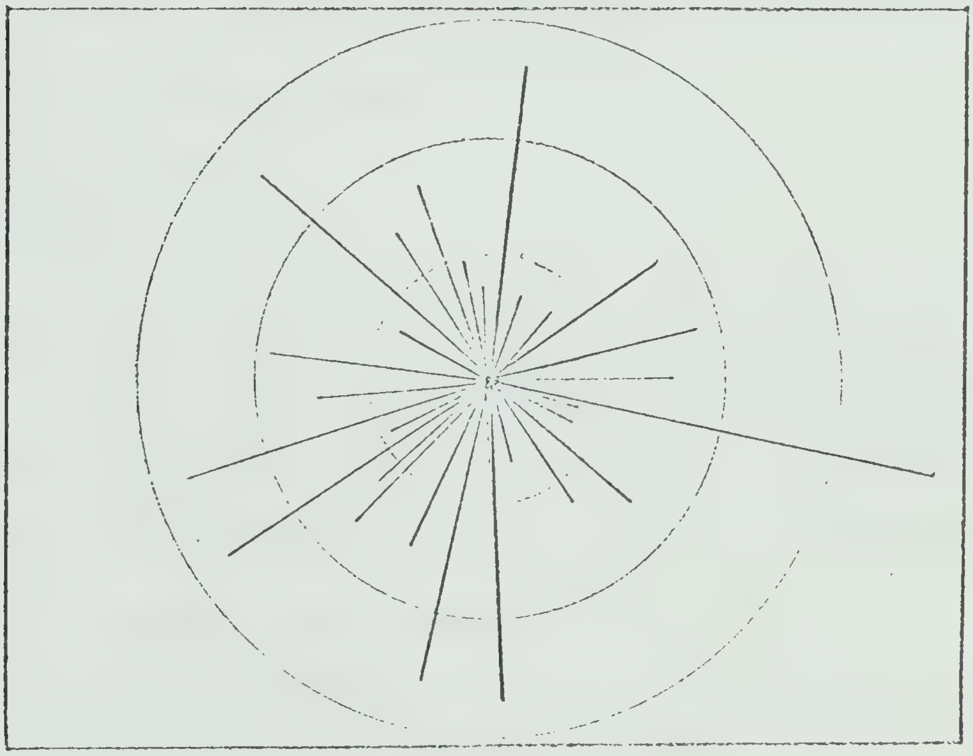
The relationship between information and distance is best appreciated through the conception of the Mean Information Field (Morrill and Pitts 1967). Information is a criterion by which one individual will tell another of an idea, or the way in which he will migrate[7]. An individual's decision will be one of a variety of possibilities. Instead of moving from place i to place j, he might move to place k. Hägerstrand (1957) assumed that all individual decisions are unique, yet have an underlying empirical regularity based on the inverse-distance relationship. This empirical regularity can be derived if individual behaviour is aggregated. The individual information or interaction field can be traced (Figure 3a), and a Mean Information Field created by collecting the origins to a single point (Figure 3b). Each information field, although unique, represents a chance variant of a general theme. If the variants are collected together, then the underlying regularity becomes apparent. In a given region, observation of the behaviour of a large number of individuals can lead to a frequency count of their movements or contacts in certain distances or directions.

---

[7]Information, here, is being viewed in a limited sense. For a more formal statement of the Theory of Information, see Weaver (1966) and Rapaport (1966).



a. Individual Information Field



b. Mean Information Field

FIGURE 3

MEAN INFORMATION FIELD

Having derived the Mean Information Field from individual behaviour, the procedure can be reversed, and the moves or contacts of a typical individual can be predicted in a deterministic manner, such that one quarter of all migrants may be said to move a distance less than ten miles. This deterministic statement may be converted to a probabilistic expression, such as the distance-decay function. In the simulation, the behaviour of typical migrants can be varied by the utilization of a random number selection procedure. The latter permits irrational behaviour, so that deviations from the systematic regularity, in keeping with the real world situation, can be introduced.

In actual simulations, one information field for all regions in the study area is uncommon. Conventionally, a different field for each region is employed, since adjustments are normally made for varying population and area sizes.

8. Distance-Decay Functions

In any model of migration or interaction, the empirical regularity of intensity of contacts varying inversely with distance can be summed up by fitting a function (Morrill 1963 a). If a graph is used, then this takes the form of fitting a curve to the observations. These functions can be treated in two ways,

a) $f(d)$, the probability or frequency of migration to a band area, which is distance d from some origin;

b) $p(d)$, the probability of migration between unit areas which are distance d apart.

The form chosen depends on the results required or the technique used. In considering interregional migration, the $p(d)$ form seems appropriate.

The technique of fitting a function to tabulated data is treated in detail in Appendix A. Reference to the graph (Figure 4) indicates that the number of contacts per unit area falls off with increasing distance away from a point of origin.

A variety of functions has been proposed by Morrill (1963a) and Marble and Nystuen (1963). The more complex Gamma, Gaussian, Fourier, and Quadratic varieties provide certain theoretical advantages, which are often outweighed by the difficulty of fitting them. The easiest, the log-normal, exponential, and the Pareto and Pareto-exponential forms are most frequently used. They provide fairly good results and can be given a linear form by log transformation, so that the parameters can be estimated by least-squares techniques (Appendix A).

Although, the fitting of a function is essentially practical, each function carries its own assumptions about the underlying data. For instance, the normal distribution would appear adequate if people travelled completely randomly over all distances and in all directions. In practice, there is a tendency for moves to be concentrated over short and very long distances. The log-normal form,

$$p(d) = ce^{-b(\log d)^2} \tag{2.3}$$

and the exponential

$$p(d) = ce^{-bd} \tag{2.4}$$

where d is the distance between points, c and b are empirically derived constants, and e is the base of the natural logarithm, have been shown to underestimate close-in movements, whereas the Pareto function overestimates such movements (Malm, Olsson, and Wärneryd 1966, p. 10). Sometimes Pareto-exponential forms have been utilized,

$$p(d) = cd^{-1}e^{-bd} \tag{2.5}$$

but the most commonly used expression is the Pareto (2.6), adjusted to take area into consideration,

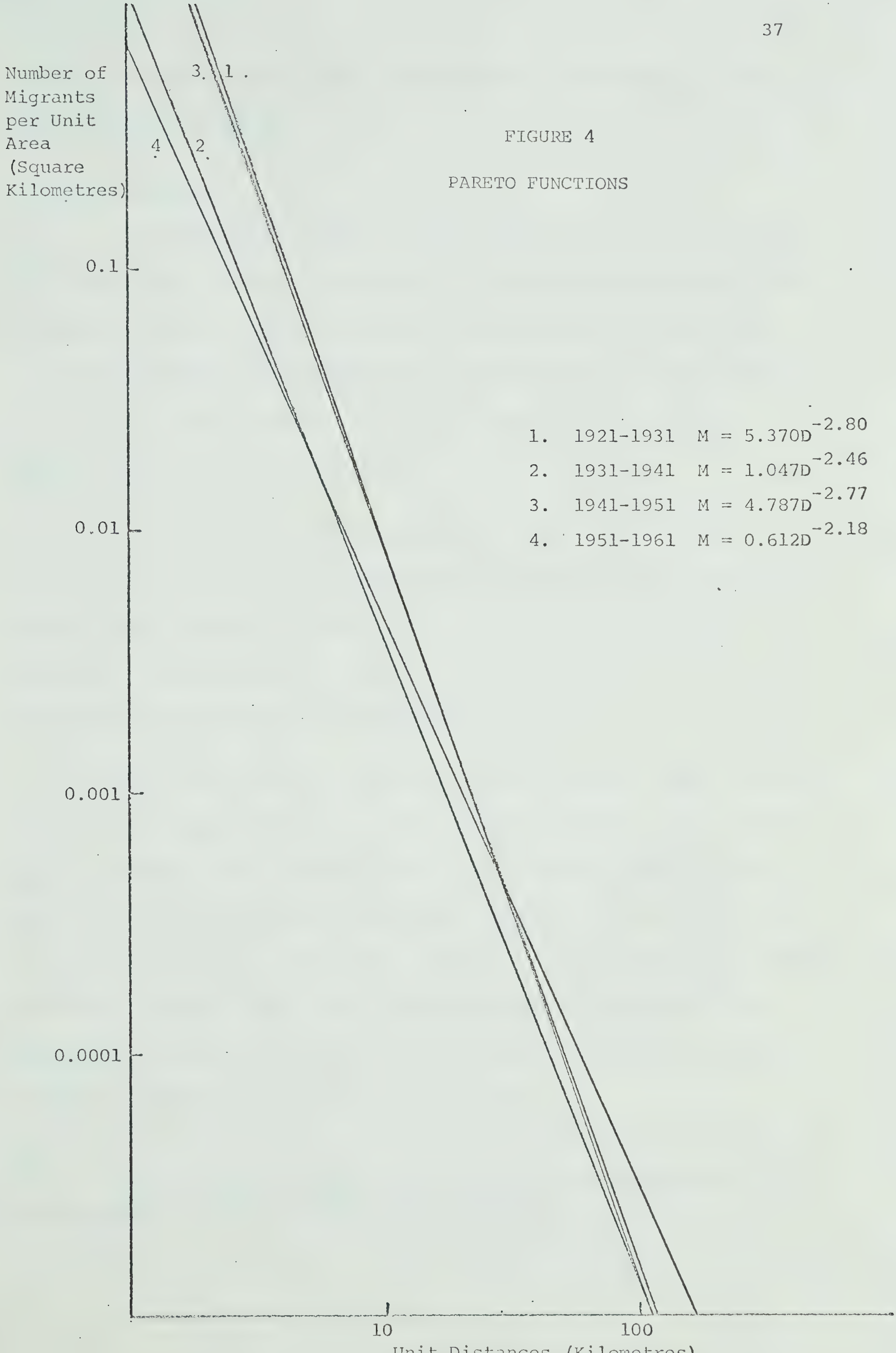
$$p(d) = \frac{c(A)}{d^b} \tag{2.6}$$

where $p(d)$ is the probability of contact between regions, d is the distance between regions' centres, A is the area of the destination region, and c and b are empirically derived constants. The value of the Pareto function has been demonstrated in many cases (Hägerstrand 1957, p. 126), and the exponent b enables results from different studies to be directly compared (Olsson 1965a, pp. 25-26). The exponent is significant in that it reflects the gradient or extent of the Mean Information Field.

The distance-decay function can be employed directly in its mathematical form in a model, or utilized in a special probability grid. If the latter technique is used, then the poor prediction of close-in contacts is eliminated by taking an empirical reading for the cell 13

FIGURE 4

PARETO FUNCTIONS



location. With the former, some modification to the function can be made at short distances[8].

9. Space Preferences

The assumption behind a distance-decay function is that there is a symmetrical fall-off in intensity of contacts away from some source. In practice, this is rarely the case. Because density of population is considered important, probabilities are weighted by the population of the source or destination area. This is a characteristic of the deterministic gravity model as well.

Other factors, including physical layout, socio-economic components of population, and communication links can disturb the symmetry (Malm, Olsson, and Wärneryd 1966, pp. 11-12). It is difficult to integrate such variables into functions, and still maintain the theoretical significance of the function.

One alternative is a suitable transformation of space to obtain a functionally based distance. In the absence of such a value, however, space preferences have been dealt with fairly simply in most models. Physical obstacles have been introduced as barriers, which reduce the intensity of contacts or the number of moves. Other attempts have created functional barriers (Yuill 1964), while multiple tellings have been employed as measures of psychological resistance, reflecting socio-economic characteristics (Pitts 1963, p. 114).

---

[8]One example of modification to the Pareto function over short distances is Pitt's (1967) MIFCAL programme.

In general, alterations to the basic probability of interaction between regions can be summed up by (Pitts 1963, p. 114) as

$$E_{ji} = \frac{P_i \cdot H_i}{\sum_{i=1}^{n} P_i \cdot H_i} \cdot B \qquad (2.7)$$

where E_{ji} is the modified interaction between regions j and i, P_i is the smoothed probability between j and i, H_i is the population of region i, and B is a barrier (equal to 0.0 if the barrier is absolute, and 1.0 if non-operative).

Marble and Nystuen (1963,p. 108) have shown, that in the case of household travel movements in Cedar Rapids, there is a strong directional bias towards the Central Business District. If this is the case in other studies, then information fields should be modified to include directional bias.

Conventionally, the diffusional process is seen as a wave moving out from a centre(s). The idea of filtering down an urban hierarchy is introduced as a modification (Hudson 1969b). To some extent, this can be envisaged within the migration process as well.

The Construction of the Monte Carlo Model

The proposed procedure is illustrated by means of the flow chart (Figure 5). The inputs include the adjusted population Ap of each region at time t , the number of potential migrants mj from each region during t to $t + \theta$, the area of each region, a subroutine

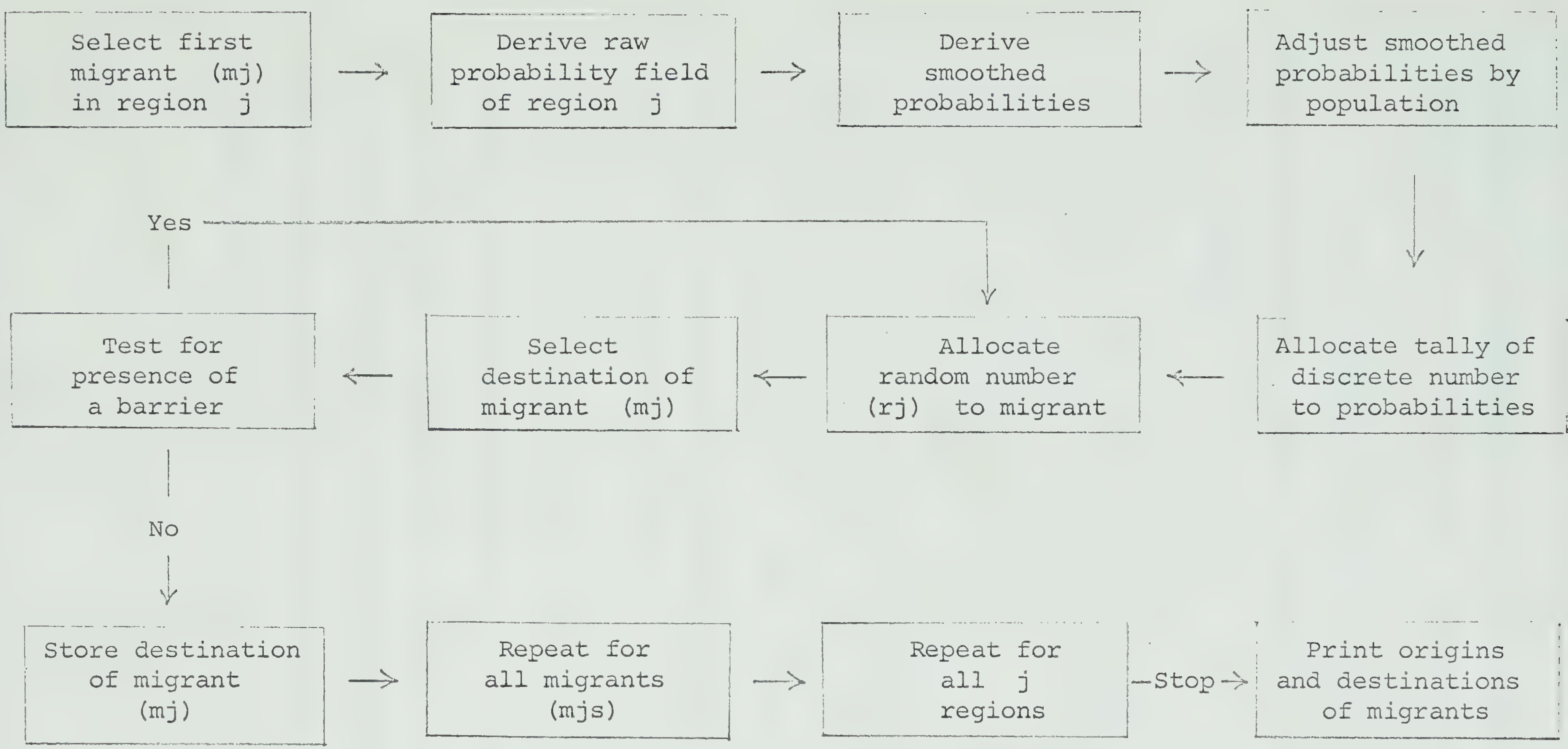


FIGURE 5

FLOW CHART OF MONTE CARLO MODEL

which calculates the distance between regions, effective barriers, the number of generations and simulations required, and a routine which averages out the results of a number of simulations.

The preceding section has treated in detail most of the features of the model. However, discussion of some of these features should be expanded.

1. The Adjusted Population: Operational Mode

The necessity of treating time in discrete periods and segregation of the elements of population change has already been mentioned. Since the study is essentially concerned with internal or local redistribution of population, it seems reasonable to accept natural increase and external migration as contraints to the model given from outside (Morrill 1965b,p. 101). How are such constraints fitted into the model? The author wishes to emphasize again that no operational procedure can be completely accurate.

For each region in each time period, t to $t + \theta$, the estimated natural increase is added to the population at time t. This new value is then adjusted to take external movements during the time period into consideration. The adjusted population Ap becomes the population of region j during time period, t to $t + \theta$, from which a certain number of migrants mj will move. Thus,

$$Ap = (P_t+N_\theta) \pm E_\theta \tag{2.8}$$

where Ap is the adjusted population, P_t is the population at the

beginning of the period, N_θ is the natural increase over the period, and E_θ is the net external gain or loss.

Migration data calculated from Census and Vital Statistics are net figures. The net figure for each region contains both external and internal components. These were segregated by using the sample as a microcosm of the settlement process in the Peace. A history of population growth and distribution was created by using the sample individuals. From this, the external and internal components of migration in each time period, for each region were estimated. Thus, actual population figures indicating a net migration total are equivalent to (External Gain + Internal Gain) - (External Loss + Internal Loss). If the relative percentages of movements are known from the sample, then the net external gain or loss can be estimated. This is equivalent to E_θ .

No account is taken of internal moves because they are simulated. The new population of a region Np at the end of a simulated time period is equal to the old, Ap, plus gain or loss of migrants mj .

2. Potential Number of Migrants

Assuming that a standard or adjusted population for each region in each time period can be derived, what proportion mj of that total can be expected to move internally (including movement within that region)? Obviously, such a figure has theoretical implications and should reflect the relative attractiveness of region j in relation to all other regions i , taking into consideration factors of employment opportunities, economic characteristics, age and sex structure, and so forth (Morrill 1965b, p. 97). The implications of such a figure are noted but in this study the derivation of mj's by substantial

theoretical reasoning was not possible because of the general paucity of relevant data. Instead, the percentage of migrants moving internally from any region was estimated from the sample.

3. Probability of Interaction Between Regions

a) Raw Probabilities

Probabilities were based upon a Pareto distance-decay function. This function has generally performed well and gives added interpretation to the migration process through the exponent b. As the distance between centres of regions was considered, a $p(d)$ form was used. Expression (2.6) gives the raw probability of contact between areas. Owing to the irregular size of regions, the function was employed directly rather than through the special probability grid.

b) Smoothed Probabilities

The raw probability values were converted to smoothed probabilities, P_i, by summing all the probabilities in the migration field of j to unity, and expressing the raw probability of contact $p(ji)$ as a fraction of unity.

c) Population Correction

The third stage requires the influence of population size of regions to be incorporated. Thus,

$$E_{ji} = \frac{P_i \cdot H_i}{\sum_{i=1}^{n} P_i \cdot H_i} \qquad (2.9)$$

where E_{ji} is the adjusted probability of interaction between regions

j and i , P_i is the smoothed probability, and H_i is the population of region i .

4. Random Number Selection

The adjusted probability of movement between regions j and i is converted to a discrete tally of random numbers. This range of numbers is relative to the ranges of all probabilities in the migration field of j . The programme selects a random number rj which determines the destination of the migrant mj .

5. Barriers

The rj value is either accepted or rejected depending on the presence or absence of a barrier to movement between regions j and i . No evidence of directional bias was evident, so directional corrections did not have to be made. The actual process of settlement does show that different regions, at different times, have only sparse settlement. In part, this is due to unfavourable physical characteristics, but also to the progression of agricultural settlement. In keeping with this, a certain number of barriers to movement have been incorporated, so that the simulation remains more faithful.

The incorporation of barriers illustrates one modification to the symmetry of the distance-decay expression. Note that a barrier between regions j and i does not necessarily imply the reverse, a barrier between i and j .

6. Time

Two aspects are involved with time. In any one time period, how many generations in the simulation are required if one generation

is equivalent to all possible migrants moving once? Using the sample, the number of generations in a simulation was made equal to the average number of moves made by actual migrants in that time period.

The second aspect involves the stability and variation of probabilities through time. In each simulated time period, the raw and smoothed probabilities were considered constant for all generations. Since each new generation results in a new population distribution Np , population adjustment to the basic probabilities was made for every succeeding generation.

In practice, the time aspect could be ignored, since the average number of moves of migrants was less than two for all time periods.

7. "Averaged" Solution

To approach the omega distribution, the results of fifty simulation runs are averaged out. A sub-routine calculates the mean population of each region in each generation.

The general layout of the programme is indicated by the flow chart (Figure 5). Detailed input characteristics are considered in Chapter Four. The programme computes for each generation:

a) the number of migrants from each region to every other region

$$\sum_{i}^{n} \sum_{j}^{n} (mji) \qquad (2.10)$$

in the form of a 26 × 26 table;

b) the number of migrants moving from each j source to all i destinations;

$$\sum_{i=1}^{n} mji \tag{2.11}$$

c) the number of migrants moving to each i destination from all j sources;

$$\sum_{j=1}^{n} mji \tag{2.12}$$

d) the net gain or loss of migrants in a region k

$$gk = \sum_{j}^{n} mji - \sum_{i}^{n} mji \tag{2.13}$$

where gk is the migrant gain or loss, $\sum_{j}^{n} mji$ is the net gain, and $\sum_{i}^{n} mji$ is the net loss;

e) the new population distribution at the end of a generation after all migrants have been redistributed

$$Np = Ap \pm gk$$

where Np is the new population, Ap the old, and gk is the net gain or loss of migrants;

f) the "averaged" population of each region after a series of simulations.

The results of the Monte Carlo model are given in Chapter Four. Evaluation and analysis are presented in Chapter Six.

CHAPTER THREE

MARKOV CHAIN MODELS

Not all stochastic models of migration, diffusion, and similar processes are simulation approaches based on sampling techniques. A variety of analytic models has been proposed, such as the information flow model of migration (Berry and Schwind 1969), epidemiology (Brown 1968a), and random and biased net constructs (Brown 1968a). Although these analytic models are applicable, determination of the parameters is often a basic problem. Brown (1968a, p. 75) described the process of finding parameters for his suggested distance-biased net model of diffusion as a series of hit-and-miss trials.

The estimation of parameters often involves circular reasoning, in which the parameters of the theoretical distribution are derived from the observations which the theoretical model should generate. Such criticism has been levelled against gravity interaction models. Since the weights of the masses and the distance are derived directly from observations, it is difficult to provide such weights with theoretical significance. Conventional use of distance-decay functions in Monte Carlo methods is also not exempt from this criticism. Where such functions are employed, the parameters are derived from the set of data under consideration. The nature of observations directly influences these parameters. Hence, different exponent values in the Pareto expression have been derived in various studies.

The weakness of many conventional analytic and simulation models has encouraged geographers to seek new methods. The possibility of

applying Markov chain models to problems in geographic research has been tentatively explored (Olsson and Gale 1968), though applications are still relatively infrequent. In the context of interregional migration, Rogers (1966b) noted that the use of a Markov Chain model permitted the population process to be separated from the population in question which is undergoing the process. Subsequently, Compton (1969) has used a similar model to describe the effect of a migration system upon a population structure, assuming fertility and mortality factors to be held constant. In this study, a finite Markov chain model is applied to interregional migration in the Peace River District[1].

> A finite Markov chain is a stochastic process which moves through a finite number of states, and for which the probability of entering a certain state depends only on the last state occupied (Kemeny and Snell 1960, p. 207).

A Markov chain consists of n discrete states of being[2]. The probability of some phenomena being in a state at time t (or after t discrete time intervals) is given by $s_{it} = (s_{1t}, s_{2t}, \ldots, s_{nt})$, where s_{it} denotes the probability of being in state i at time t. Since, phenomena can be distributed only among the n discrete states at time t, then

$$\sum_{i=1}^{n} s_{it} = 1 \tag{3.1}$$

---

[1]Most of the material on Markov chains has been derived from the work of Kemeny. These include Kemeny and Snell (1960), Kemeny, Snell, and Thompson (1956), and Kemeny (1959); note also Bharucha-Reid (1960).

[2]A state could be regarded as a class, a region, or any suitable category.

The probability of moving from state i to state j within one discrete time period is denoted by p_{ij} . These transition probabilities form a transition matrix, P . Each row of P is a probability vector, since phenomena in state i at time t can move only to each of the remaining $n - 1$ states, or remain in state i by time $t + 1$. Hence,

$$\sum_{j=1}^{n} p_{ij} = 1 \tag{3.2}$$

The matrix P , composed of transition probabilities, can be written as

$$P = \begin{array}{cc} & \begin{array}{cc} S_i & S_j \end{array} \\ \begin{array}{c} S_i \\ \\ S_j \end{array} & \begin{bmatrix} p_{ii} & p_{ij} \\ & \\ p_{ji} & p_{jj} \end{bmatrix} \end{array} \tag{3.3}$$

where the states are S_i and S_j , and p_{ii} , p_{ij} and so forth are the transition probabilities.

The fundamental Markovian property is that a particular outcome, occurring in a certain state at a certain time, is directly dependent upon the preceding outcome in the previous time period. Let x_t be the state that a phenomenon is in at time t , where x_t is a random variable, whose chance of occurrence is denoted by a probability

measure[3]. The conditional probabilities can be defined (Hammersley and Handscomb 1964, p. 113),

$$P(x_t=S_j \mid x_{t-1}=S_{i_{t-1}}, \ldots, x_t=S_{i_2}, x_1=S_{ij})$$
$$= P(x_t=S_j \mid x_{t-1}=S_{i_{t-1}}) \tag{3.4}$$

and

$$p_{ij} = p(S_i \rightarrow S_j) = P(x_t=S_j \mid x_{t-1}=S_i) \tag{3.5}$$

A characteristic of some types of Markov chains is the derivation of a state of statistical equilibrium. Interpretation of this equilibrium is valuable in geographic problems. The derivation of equilibrium solutions will be treated in detail in a later section.

Markov chains are not entirely exempt from the criticism of circular reasoning. Compton (1969) estimated transition matrices from population data. Clark (1965) also derived matrices from the data under consideration. Even where matrices are developed through the use of distance-decay functions or gravity interaction expressions, there is still the problem of giving meaning to the parameters of these distributions. Consequently, the derivation of a transition matrix

---

[3]A random event has a chance of occurrence. Its probability is a measure of that chance, ranging from complete certainty that it will occur (unity), to complete certainty that it will not occur (zero).

remains an important issue in the creation of Markov chain models.

Markov chain models do not easily permit the introduction of other significant variables. Olsson and Gale (1968) proposed the use of q-dimensional matrices, though this raises the question of how to define appropriate multiplication rules. However, Rogers (1966b) has suggested methods by which age and sex structure can be incorporated into matrix operation.

Despite these limitations, many problems can be approached by a model, which explicitly states that a particular event is directly dependent upon the preceding event. Golledge (1967) suggested that the Markovian property is extremely relevant in dealing with the market-decision process. Such a process is characterized by a search, in which a learning sequence gradually produces a habitual response pattern. Each decision is influenced by the preceding decision. The habitual response can be regarded as an equilibrium situation.

Consequently, use of Markov chain models is steadily increasing in the Social Sciences. Among economists, Hampton (1968) developed a transition matrix showing the probability of movement of industry between regions of New Zealand. Smith (1961) utilized the technique to consider interregional trade. Harris (1968) proposed the model as a means to deal with the question of urban growth. The model could be used to indicate direction and timing of city expansion. Among geographers, Brown (1963) has developed a Markovian model of diffusion while Marble (1964) derived equilibrium solutions of trip movement structure. Rogers (1966a) saw equilibrium distributions of population as targets or guides for government policy on interregional

migration.

There is a need for further exploitation of the model. The gaps left by simulation techniques are considerable. Markov chain models may be one means of filling them. While this study specifically deals with migration, the implication is that the model could be extended to other forms of spatial behaviour or interaction. Before turning to the construction of a Markov chain model, some relevant aspects of these models will be illustrated.

Regular Markov Chains

Regular chains are an important class of Markov chains. "A Markov chain is called a regular chain if some power of the transition matrix has only positive elements" (Kemeny 1959, p. 391). A regular chain is ergodic, though an ergodic chain need not be regular. "A Markov chain is called an ergodic chain if it is possible to go from every state to every other state" (Kemeny 1959, p. 394).

To observe a Markov chain process, the transition matrix and the starting state are specified. The process moves through a set of states successively. If the states and the transition matrix can be properly defined, then the behaviour of the system through time can be computed (Kemeny and Snell 1960, p. 208). The possible sequence of movements between states, and the probability of being in any state after so many steps or discrete time periods can be derived. A simple example will clarify this description.

Suppose, there are two states, S_1 and S_2. The initial

probability vector is $p^{(o)} = (p_1^{(o)}, p_2^{(o)})$, where $p_1^{(o)}$ denotes the probability of phenomena being in S_1 at this initial stage, and $p_2^{(o)}$ denotes the probability of being in S_2. Let the transition matrix be

$$P = \begin{array}{c c} & \begin{array}{cc} S_1 & S_2 \end{array} \\ \begin{array}{c} S_1 \\ \\ S_2 \end{array} & \begin{bmatrix} a & 1-a \\ & \\ 1-b & b \end{bmatrix} \end{array}$$

Knowing the initial distribution of phenomena, what will the distribution be after one step or discrete time period? These probabilities can be estimated by a simple procedure of multiplication of conditional probabilities (Kemeny 1959, p. 390). The probability $p_1^{(1)}$, of being in S_1 after one step is

$$p_1^{(1)} = p_1^{(o)} \cdot a + p_2^{(o)} \cdot 1 - b$$

and the probability of being in S_2 after one step is

$$p_2^{(1)} = p_1^{(o)} \cdot 1 - a + p_2^{(o)} \cdot b$$

Thus, the probability of being in state i after n steps is the sum of the probabilities of being at each of the possible states at $n - 1$ steps, and then moving to state i on the nth step,

$$\begin{aligned} p_1^{(n)} &= p_1^{(n-1)} \cdot p_{11} + p_2^{(n-1)} \cdot p_{21} + \cdots + p_m^{(n-1)} \cdot p_{m1} \\ &\vdots \\ p_m^{(n)} &= p_1^{(n-1)} \cdot p_{1m} + p_2^{(n-1)} \cdot p_{2m} + \cdots + p_m^{(n-1)} \cdot p_{mm} \end{aligned} \tag{3.6}$$

Expression (3.6) can be written as

$$p^{(n)} = p^{(n-1)} \cdot P \tag{3.7}$$

which satisfies the above equations. The multiplication of the matrix P, by a probability vector p, considerably reduces the calculations required to estimate the probability of being in any state after n steps. The numerical example given in Appendix B will further illustrate these comments.

It can also be shown (Kemeny 1959, p. 402), that given expression (3.7), then

$$p^{(n)} = p^{(o)} \cdot P^{n} \tag{3.8}$$

and

$$P^{n} = P^{n-1} \cdot P = P \cdot P \cdot P \cdot \ \ldots \ P \quad (n \ \text{ factors}) \tag{3.9}$$

If the vector $p^{(o)}$, of initial probabilities is multiplied by the nth power of P, then the vector $p^{(n)}$ is obtained, the components of which give the probability of being in state j after n steps. Thus, $p_{ij}^{(n)}$, the ijth entry of matrix P^{n}, states the probability of the process being in state j after n steps, if it started in state i.

Matrix multiplication provides a convenient way of obtaining the desired probabilities (Appendix B). The probability of moving from

state i to state j is always the same. The constant or stationary nature of transition probabilities is a feature of Markov chains.

Equilibrium Solutions

The attainment of a state of equilibrium is one of the most important characteristics of regular Markov chains. The vector $p^{(n)}$ is obtained from the vector $p^{(n-1)}$ by multiplying it by the matrix P. The vector $p^{(n-1)}$ is obtained from $p^{(n-2)}$ is a similar fashion, and so on (Kemeny 1959, p. 402). The vectors obtained by this linear transformation can be indicated by

$$p^{(o)} \rightarrow p^{(1)} \rightarrow p^{(2)} \rightarrow \ldots\ p^{(n-1)} \rightarrow p^{(n)} \rightarrow \tag{3.10}$$

The vector $p^{(o)}$ is sent on to the vector $p^{(1)}$, $p^{(1)}$ on to $p^{(2)}$, and so forth. It sometimes happens that there is a probability vector t, which is sent by the transformation of P on to itself, that is

$$t = t \cdot P \tag{3.11}$$

This vector t is termed the fixed point of the transformation P. It follows that the powers P^n approach a matrix T, each row of which is the same probability vector t (Kemeny 1959, p. 416).

Suppose, there is a system composed of three states of being—Rain, Nice, Snow (Kemeny 1959, pp. 384-438). Assume that a discrete time period is equivalent to one day. The probability of moving from one state to another, that is of the weather changing by days, is

given by

$$\begin{array}{c|ccc} & R & N & S \\ \hline \end{array}$$

$$P = \begin{array}{c} R \\ N \\ S \end{array} \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

That the powers P^n approach a matrix T can be seen since

$$P^8 = \begin{array}{c} \\ R \\ N \\ S \end{array} \begin{array}{c} \begin{array}{ccc} R & N & S \end{array} \\ \begin{bmatrix} 0.400 & 0.200 & 0.400 \\ 0.400 & 0.200 & 0.400 \\ 0.400 & 0.200 & 0.400 \end{bmatrix} \end{array}$$

Here, $p_{ij}^{(8)}$, denotes the probability of the process being in state j after eight steps, if it started in state i . The probability of having a nice day eight days after rain is 0.200. This is nearly the same (to three decimal places) as the probability of having a nice day eight days after snow or nice. Obviously, this is close to the equilibrium solution. Such a solution can be determined by solving a system of simultaneous equations. From (3.11),

$$(t_1,t_2,t_3) \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix} = (t_1,t_2,t_3)$$

then

$$\begin{aligned} t_1 + t_2 + t_3 &= 1 \\ \tfrac{1}{2}t_1 + \tfrac{1}{2}t_2 + \tfrac{1}{2}t_3 &= t_1 \\ \tfrac{1}{4}t_1 \qquad\quad + \tfrac{1}{4}t_3 &= t_2 \\ \tfrac{1}{4}t_1 + \tfrac{1}{2}t_2 + \tfrac{1}{2}t_3 &= t_3 \end{aligned}$$

The unique solution is $t = (0.4, 0.2, 0.4)$. The matrix T is

$$T = \begin{array}{c} \\ R \\ N \\ S \end{array} \begin{array}{c} \begin{array}{ccc} R & N & S \end{array} \\ \begin{bmatrix} 0.4 & 0.2 & 0.4 \\ 0.4 & 0.2 & 0.4 \\ 0.4 & 0.2 & 0.4 \end{bmatrix} \end{array}$$

The probability of the process ending up in state j is independent of the starting state. Irrespective of whether the process starts in R, N, or S, the equilibrium solution assumes that the probability of it ending up in R is always 0.4 , in N is 0.2 , and in S is 0.4 .

How is this statistical equilibrium to be interpreted? The unique solution above implies that it will rain 40% of the time, snow 40% , and be nice 20% . In the case of interregional migration, such a steady-state will be reached when out and in migration from each region are balanced in such a way that the system remains stable. No region will increase or decrease in size relative to the others, once this state is reached. The steady-state reflects the ideal or theoretical distribution which should emerge over time. Markov chains are dependent upon time, though once the equilibrium state is achieved, they become time independent. The equilibrium solutions can be derived without reference to the initial distribution.

The Collective Process

The individual process has been considered, in which the $p_{ij}^{(n)}$ entry denotes the probability that an individual who started in the ith state will be in the jth state after n steps. The complementary collective process is more relevant to the study of interregional migration (Kemeny and Snell 1960, p. 209). In the collective process, the ijth entry of P^n represents the fraction of people who started in the ith region and will be in the jth region after n intervals.

Assume that a transition matrix has been derived showing the probability of movement between three regions (Rogers 1966a, p. 207), such that

$$P = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 1/2 & 1/4 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

The initial distribution of population is given by $p^{(o)} = (\frac{4}{9}, \frac{11}{16}, \frac{1}{12})$, where $\frac{4}{9}$ of the total population are located in region one, and so on. In practice, it makes no difference if absolute figures are used instead of fractions. Thus, $p^{(o)} = (16.0, 11.0, 9.0)$, where $p^{(o)}$ contains the population of the regions. Using (3.7) the distribution of population after several intervals can be calculated.

$$p^{(1)} = (11.08, 13.83, 11.08)$$

$$p^{(2)} = (10.84, 14.34, 10.84)$$

From (3.11) the equilibrium solution is $t = (10.84,14.34,10.84)$. This distribution is arrived at regardless of the initial conditions and is, therefore, a steady state. It represents a balanced condition in which inputs and outputs from each region are stable. Over time, region one will come to contain 10.84 , region two 14.43, and region three 10.84. Once the state of equilibrium is reached, these amounts will not vary. If the vector t is multiplied by P , the same answer will result.

Absorbing Markov Chains

In the simple Markov chain all processes are reversible. This need not be so. Labour and capital turned into a road cannot be recreated by breaking up the road (Harvey 1967, p. 580). In regular Markov chains, the process can go from every state to every other state. Absorbing Markov chains contain at least one state from which it is impossible to leave once reached, and from every other state it is possible to go to an absorbing state, not necessarily in one step (Kemeny and Snell 1960, p. 208). An absorbing chain can be recognized by the unity values in the appropriate p_{ii} entries.

Suppose, there are five regions between which migration can take place, except that one and five cannot be left once entered, but can be reached from all other regions. A matrix P can be written (Kemeny and Snell 1960, p. 45),

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Conventionally, the matrix of an absorbing chain is written in the canonical form,

$$P = \begin{array}{c} \text{r states} \\ \text{s states} \end{array} \begin{array}{c} \begin{array}{cc} \text{r states} & \text{s states} \end{array} \\ \left[\begin{array}{c|c} I & O \\ \hline R & Q \end{array}\right] \end{array} \tag{3.12}$$

in which exist r absorbing states and s non-absorbing states. The matrix is partitioned into four sub-matrices. I is a $r \times r$ identity matrix with ones in the diagonal, O is a $r \times s$ zero matrix. R, the $s \times r$ matrix, concerns the transition from transient to ergodic states, while Q, the $s \times s$ matrix, concerns the process as long as it stays in the transient states.

> A transient state is an element of a transient set ... A transient set of states is a set in which every state can be reached from every other state and which can be left ... An ergodic state is an element of an ergodic set ... An ergodic set of states is a set in which every state can be reached from every other state, and which can not be left once it is entered (Kemeny and Snell 1960, p. 208).

The above example is written as

$$P = \begin{array}{c} \\ 1 \\ 5 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 5 & 2 & 3 & 4 \end{array} \\ \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{array}\right] \end{array}$$

From this can be derived N , the <u>fundamental matrix</u> for a given absorbing chain. The fundamental matrix provides considerable information about the absorbing chain. It can be used to state the probability that the process will end up in a given absorbing state, the average length of time for the process to reach an absorbing state, and the average number of times that the process will be in each non-absorbing state (Kemeny and Snell 1960, p. 404). Here, it is merely stated that

$$N = (I-Q)^{-1}$$

where N is the fundamental matrix, and $(I-Q)^{-1}$ is the inverse of the matrix $(I-Q)$. From above,

$$I - Q = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{bmatrix}$$

and

$$N = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{bmatrix} \end{array}$$

This section has been included to familiarize the reader with the terms and technicalities of an absorbing chain. In any realistic assessment of an interregional migration system, where natural growth and external movements are considered, an absorbing chain seems more appropriate than a regular one. Use of an absorbing chain poses certain difficulties to the construction of a model of migration.

Construction of the Markov Chain Model

The aim is to create a model, which will develop an equilibrium population distribution for the twenty-six regions in the study area in each time period. How is the affect of population and external movements on the population structure to be incorporated in the model? What methods can be used to derive the transition probabilities?

1. Population Growth

Only if population remains stable can a regular Markov chain model be used. This involves the patently unrealistic assumption that birth, death, and migration rates balance perfectly. If births and deaths are included the chain becomes absorbing, since the former can not be entered from any other state, nor can the latter be left. If death is included as a state of being, conventional techniques imply that, in the equilibrium solution, everyone will end up dead. Rogers (1966a, pp. 207-209) has proposed a way out of this dilemma. He suggested that a transition matrix P be developed with death as the nth state. The chain would be absorbing. Define a vector of births $X = (x_1, x_2, \ldots, x_{n-1})$ for each region. Initial distribution of population among regions is given by $W^{(o)} = (w_1^{(o)}, w_2^{(o)}, \ldots, w_{n-1}^{(o)})$. The

death state is ignored, except for its impact upon the transition probabilities. Remembering the canonical form of P, (3.12), and given $w^{(o)}$ and P, then

$$\overline{w}^{(o)} = w^{(o)} \cdot Q \tag{3.14}$$

where Q is the set of transient states. This gives a new population, $\overline{w}^{(o)}$, which will be smaller than $w^{(o)}$, since some people will have died. If the vector, x, is added to $\overline{w}^{(o)}$, the distribution of the population, including births, at the end of one time period os obtained, $w^{(1)}$.

$$w^{(1)} = \overline{w}^{(o)} + x = w^{(o)} \cdot Q + x \tag{3.15}$$

$$w^{(2)} = (w^{(o)} \cdot Q + x)Q + x \tag{3.16}$$

and

$$w^{(n)} = w^{(o)} \cdot Q^{n} + x(I + Q + \ldots + Q^{n-1}) \tag{3.17}$$

However, $Q^n \rightarrow 0$ as n increases, and $w^{(o)} \cdot Q^n \rightarrow 0$ also as n increases, which leaves $w^{(n)} \rightarrow x(I+Q+\ldots+Q^{n-1})$. As n increases, $I + Q + \ldots + Q^n \rightarrow (I-Q)^{-1}$, and the limiting or equilibrium distribution is given by

$$y = \lim_{n \to \infty} w^{(n)} = x(I-Q)^{-1} \tag{3.18}$$

A numerical example will illustrate the procedure (Rogers 1966a, p. 208). Consider a system of three regions, A, B, C, and a state of death, D,

$$P = \begin{array}{c|cccc} & A & B & C & D \\ \hline A & 1/4 & 1/4 & 1/4 & 1/4 \\ B & 1/5 & 2/5 & 1/5 & 1/5 \\ C & 1/4 & 1/4 & 1/4 & 1/4 \\ D & 0 & 0 & 0 & 1 \end{array}$$

then

$$(I-Q)^{-1} = \begin{array}{c|ccc} & A & B & C \\ \hline A & 2.0 & 1.25 & 1.0 \\ B & 1.0 & 2.50 & 1.0 \\ C & 1.0 & 1.25 & 2.0 \end{array}$$

The vector x is $(3,2,1)$ and, using expression (3.18),

$$y = (3,2,1) \cdot \begin{pmatrix} 2.0 & 1.25 & 1.0 \\ 1.0 & 2.50 & 1.0 \\ 1.0 & 1.25 & 2.0 \end{pmatrix}$$

$$= (9,10,7)$$

Consequently, in the equilibrium distribution, regions A, B, and C will contain 9, 10, and 7, respectively.

The above procedure in modified form is followed in this study.

2. External Migration

Roger's model considered only local migration and natural growth. No account is taken of movements into and out of the entire area. A reasonable approach would adjust the x vector to include external migration as well as births, while the nth state, D, could incorporate

death rates[4].

3. The Transition Matrix

Conceivably, a matrix could be derived by reference to the movement of sample individuals. This approach was rejected on two grounds. Inspection of the results indicated that the pattern of movement in the Peace is such, that not all regions sent migrants to every other. Consequently, many cells in the 26×26 matrix were left blank. A transition matrix should be ideally based on some theoretical criterion. Accordingly, a simple gravity interaction model was used to derive the transition probabilities,

$$I_{ij} = g \frac{P_j}{D_{ij}^b} \Big/ \sum_{h=1}^{n} \frac{P_h}{D_{ih}^b} \qquad (3.19)$$

where I_{ij} is the interaction between regions i and j, D_{ij} is the distance between i and j, P_j is the population of j, P_h is the population of region h where $h = (1,2,\ldots,n)$. D_{ij} is the distance between h and i, and b and g are empirically derived constants.

Obviously, the parameters of the gravity model are open to the criticism of circular reasoning, but at least some effort is made to provide the transition matrix with theoretical significance. The distance variable in (3.19) is straight-line distance between the

---

[4]This vector X almost always contained positive elements. Where it did not, the X value was considered zero, and the net population loss was added into the death rate.

the centres of regions[5]. The masses were taken as the population of regions weighted by area[6]. In view of the controversy over which exponent b values should be used (Olsson 1965b, pp. 57-63), the values used in the model are derived from the Pareto expressions of the Monte Carlo model. The constant g was similarly derived.

For each row, a value from the gravity model was obtained for each p_{ij}. There are two exceptions. The D value was obtained from death rates, while the p_{ii} values were estimated from sample data. The cell values for each row were then normalized.

4. Stability of Transition Probabilities

Transition probabilities in the Markov chain process are stationary. They do not vary within each time period. Feedback and changing inputs will lead to a disturbance of such probabilities. Accordingly, a different transition matrix was derived for each of the four time periods.

The results of the Markov chain model for each time period are set out in Chapter Four. Detailed analysis is given in Chapter Six.

---

[5]Other conceivable measures could have been used, but apparently offer little more advantage (Olsson 1965b, p. 57).

[6]In effect, the population per square kilometer was used.

CHAPTER FOUR

EXPERIMENT

An experiment brings the models and observations together in this chapter. Discussion of the real world situation in the Peace is followed by presentations of detailed inputs to the models, and of the results of the experiment. Evaluation of the models and the experiments is reserved for the following chapter.

The Peace River Study Area

1. General Background

The broadest definition of the Peace District covers an immense area in the Canadian north-west. Extending west to east from Hudson Hope in British Columbia to Lesser Slave Lake in Alberta, the region stretches north from the Wapiti River to Fort Vermillion which is close to the great Athabasca Lake (Kitto 1930, p. 1). The study area, for practical reasons, is much smaller (47,350 sq. kms.), and embodies the core of the Peace District (Figure 1).

Basically, the Peace District, a plateau with a prevailing elevation of 2000' to 3000', is composed of glacial, fluvial, and lacustrine deposits laid down on the horizontally bedded limestones of the Albertan Syncline. Relief differences (Figure 6) are due either to the deep incision of rivers or to prominent uplands, which are remnant till plains 300' to 1200' above the plateau (Tracie 1967, p. 7).

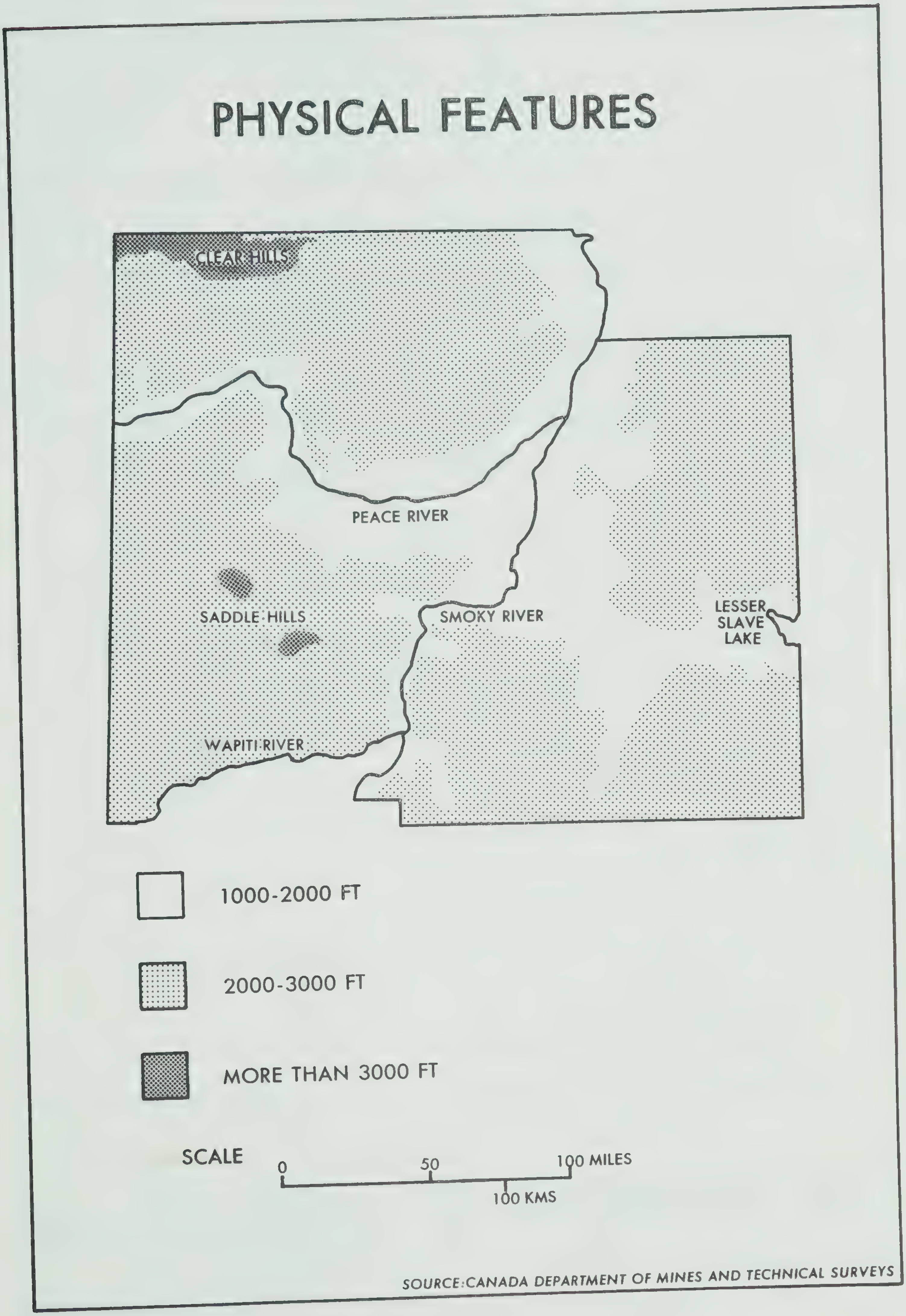

PHYSICAL FEATURES
CLEAR HILLS
PEACE RIVER
SADDLE HILLS
SMOKY RIVER
LESSER SLAVE LAKE
WAPITI RIVER
1000-2000 FT
2000-3000 FT
MORE THAN 3000 FT
SCALE
0
50
100 MILES
100 KMS
SOURCE: CANADA DEPARTMENT OF MINES AND TECHNICAL SURVEYS

FIGURE 6

Despite a northern continental location between latitudes 55° N and 57° N, the Peace region is well suited to the growing of cool-season crops (Carder 1965, p. 5). While late spring and early fall frosts can cause crop damage, and summer droughts occasionally reduce yields, the area has a number of distinct advantages. Precipitation values, although low (yearly mean 17.54" , including snowfall), are largest in July (mean 2.44") and are well distributed throughout the rest of the growing season. Moderate temperatures during summer (July mean 60°F) help keep evapo-transpiration low, and are adequate for growth. The long summer days overcome the disadvantage of the short growing season[1].

Killing frosts were once believed to be a major deterrent to agriculture, but agricultural expansion seems to have been conditioned by soil rather than frost (Tracie 1967, p. 7). Parkland areas have been the most favoured areas for settlement. Such areas are stretches of aspen parkland located on relatively level terrain, and developed on better soils such as degraded Black Earths (Tracie 1967, p. 8). The Grande Prairie (Buffalo Plains) is the largest of the parkland areas. Others are near Spirit River, Pouce Coupé, and to the north of the Peace River (Figures 7 and 8).

Prior to the early part of the 20th century, European interest in the area was confined to an almost legendary group of traders, missionaries and prospectors[2]. The isolation of the Peace prevented

---

[1]The temperature and rainfall figures are derived from Carder's Table 1 (1965, p. 6). They are based on data collected for the years 1916 to 1960 at Beaverlodge.

[2]A number of historical texts are available. The most readable of these is MacGregor's work (1952).

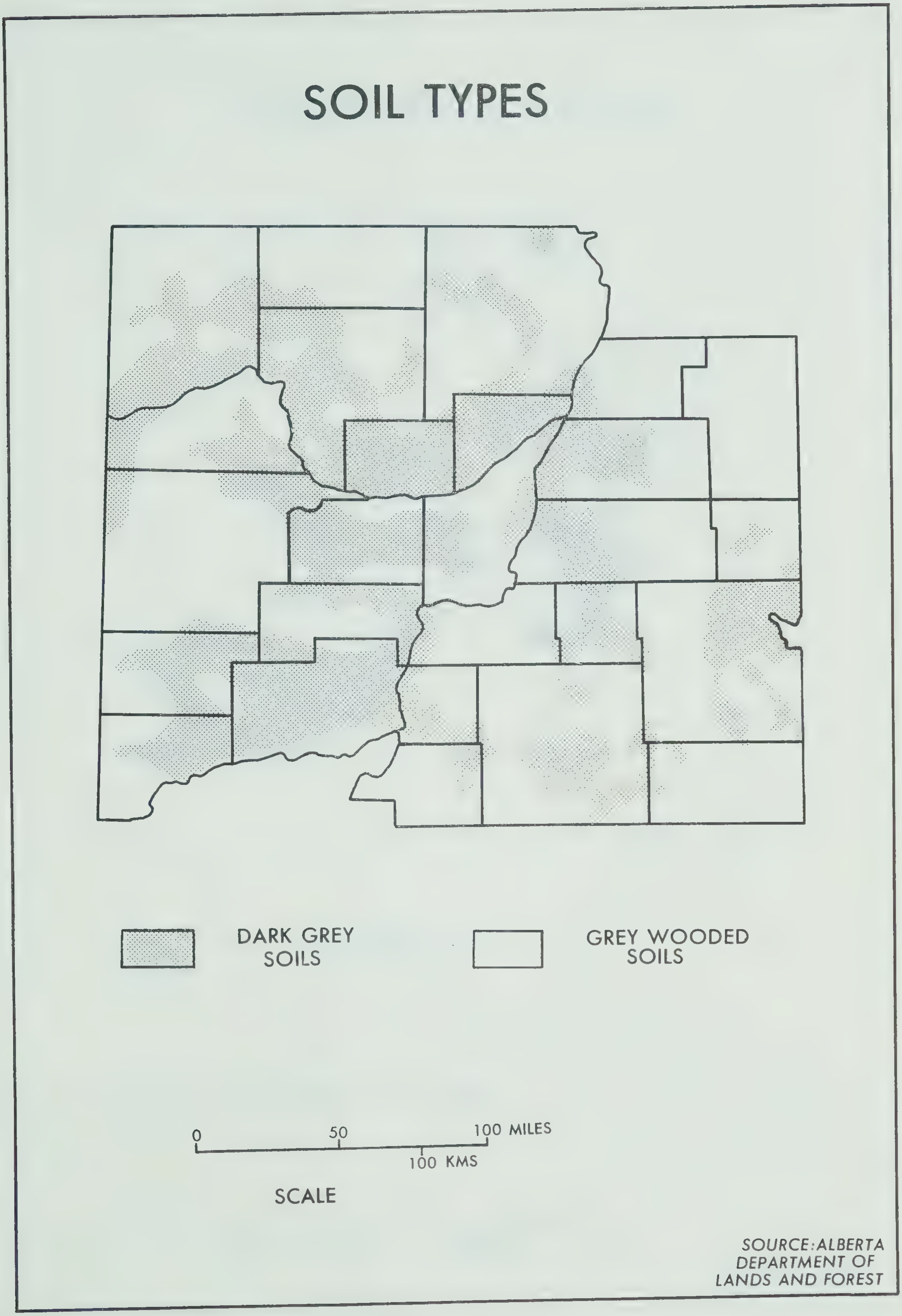
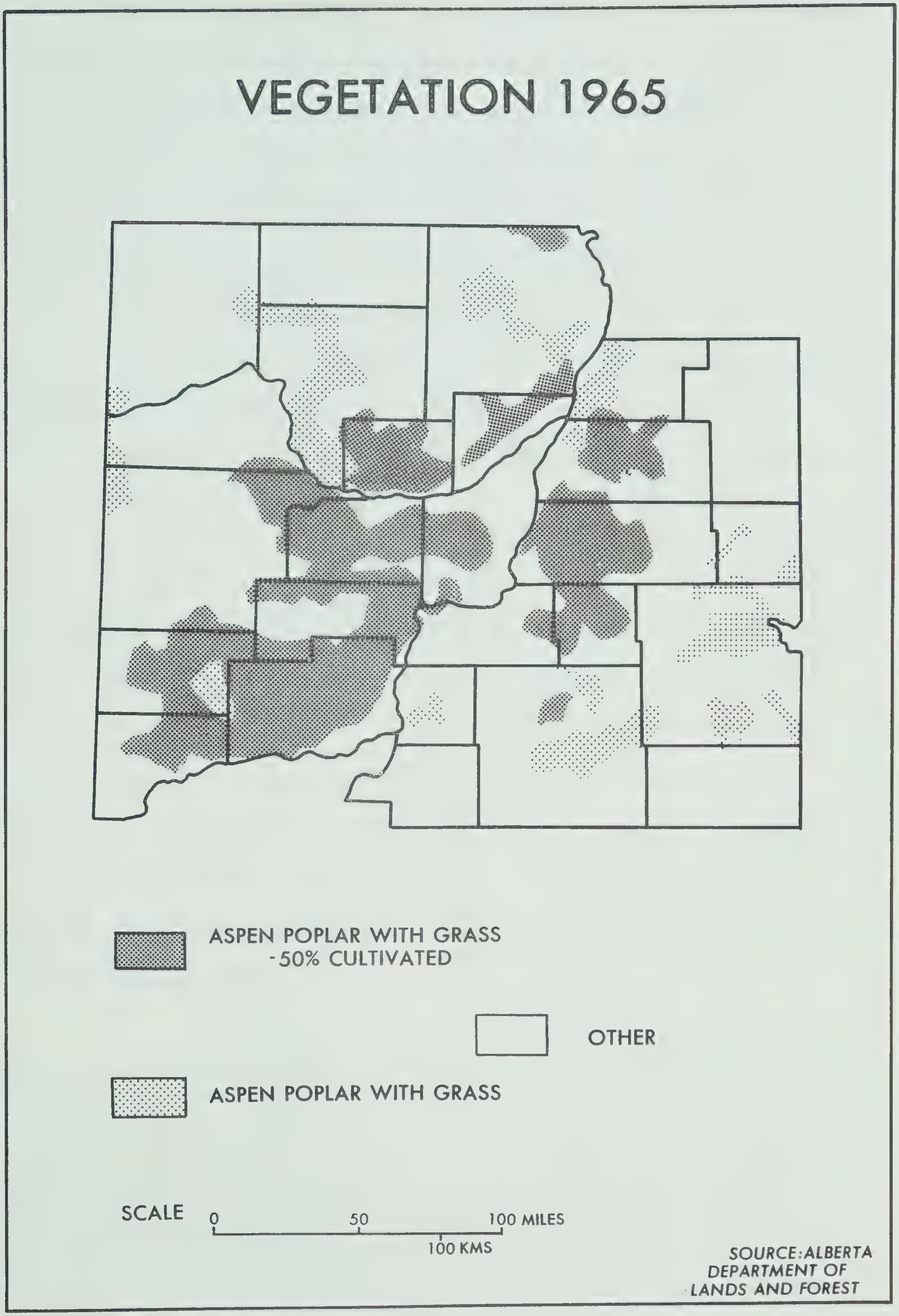

SOIL TYPES
DARK GREY SOILS
GREY WOODED SOILS
0
50
100 MILES
100 KMS
SCALE
SOURCE: ALBERTA DEPARTMENT OF LANDS AND FOREST

FIGURE 7



VEGETATION 1965
ASPEN POPLAR WITH GRASS
-50% CULTIVATED
OTHER
ASPEN POPLAR WITH GRASS
SCALE
0
50
100 MILES
100 KMS
SOURCE: ALBERTA
DEPARTMENT OF
LANDS AND FOREST

FIGURE 8

any wide spread of settlement. This isolation was doomed in the great westward movement of people after the 1870's, when the first trans-continental railway was completed. A railway did not reach the town of Peace River until 1916, but its impact was felt long before that (Willis 1966, p. 19). The demand for wheat from industrialised North-West Europe was high, and railroads provided the means by which people could be brought in and the wheat taken out. As more and more land was taken up by settlers, new areas were sought. Various reports had been circulating about the agricultural potential of the Peace. These rumours were confirmed by a number of scientific travellers (MacGregor 1952, p. 28).

From 1905 onwards, settlers began arriving without waiting for the railroad. The first were ranchers, but these were soon superseded by grain farmers. Clusters of settlement developed, including Grande Prairie and Spirit River. Communication were improved, first by the Edson Trail, which precipitated the area's first land boom in 1911, and then by the railway which reached Grande Prairie in 1916. Although, the effects of the railway were overshadowed by World War I, access was provided to a stable market, and settlers could now come in more easily (Willis 1966, p. 19).

The population of the study area (Table 4.1) has increased steadily from 1921 (16,050) to 1961 (59,256). There were fluctuations in population after both World Wars, due to liberalised homesteading schemes for war veterans. The drought years of the late 1920's in Alberta and Saskatchewan encouraged many settlers to try their luck in The Peace. In contrast, there were out movements during the war years, and many

TABLE 4.1 - AREA AND POPULATION CHARACTERISTIC OF STUDY REGIONS, 1921-1961

Region	Centre				Area	Population				
	Latitude		Longitude		(In square					
	(degrees		and minutes)		kilometres)	1921	1931	1941	1951	1961
1	116	27	54	59	1669	0	24	28	28	128
2	117	23	55	07	3339	148	475	846	1217	2522
3	118	07	54	59	1041	2	25	93	55	61
4	118	04	55	15	770	104	414	730	851	1074
5	116	28	55	22	3297	1322	2823	3159	4061	4636
6	117	10	55	30	835	77	503	601	815	942
7	117	47	55	29	1156	0	24	114	90	127
8	116	15	55	46	835	72	141	204	202	385
9	117	03	55	46	1959	1156	3309	3874	4954	4825
10	116	20	56	10	1850	0	10	0	10	0
11	116	59	56	02	1706	251	1106	1249	1662	1994
12	116	56	56	18	1096	1233	1341	1453	2195	3706
13	118	44	55	15	2234	4030	7133	7268	8229	13549
14	119	40	55	03	1420	852	2026	2112	1614	1321
15	119	35	55	20	1489	1271	3577	3745	3634	3793
16	118	40	55	32	1357	354	1385	1793	1577	1362
17	117	49	55	49	1600	430	710	1050	1609	1766
18	118	32	55	46	1360	1260	2413	2650	3233	3585
19	119	27	55	45	3838	290	1688	2046	2120	1773
20	119	24	56	07	1700	0	32	38	99	317
21	117	45	56	08	897	1337	2127	2378	2593	3302
22	118	19	56	02	1004	1456	2524	2758	2960	3423
23	118	37	56	18	2474	46	1525	2069	2300	2273
24	117	40	56	30	3529	359	1562	1844	1858	1570
25	118	38	56	38	1669	0	24	80	97	84
26	119	35	56	29	3226	0	151	357	415	736
Total					47350	16050	37072	42539	48478	59254

Source: Canada, Dominion Bureau of Statistics

people left after the bad harvests of the early 1920's (Tracie 1967, p. 23).

In recent times, the Peace has become more integrated into the Western Canadian economy. Partly, this stems from the exploitation of natural oil and gas, partly, from improved communications (Vanderhill 1963). The construction of the Alaskan (1942) and MacKenzie (1948) Highways and the Great Slave Lake Railway have linked the region to resource developments in the North-West. In 1955 a direct paved highway to Edmonton via Valleyview was completed.

Homesteading policies have encouraged the development of the area with varying degrees of success. Renewed interest in the late 1950's and early 1960's led to the opening up of new areas for agriculture, including south of the Wapiti River. Yet, while occupied land has increased, the number of farms has declined. This seems to be more the result of the enlargement and consolidation of land holdings, rather than the activities of new settlers (Vanderhill 1963, p. 38).

2. The Settlement Process

The primary aim of this thesis is to investigate the relative efficiency of the Monte Carlo and Markov chain models, not to examine the process of agricultural settlement. However, some idea of the gradual spread of settlement and transportation in the Peace can be gained by examining the series of maps from Figure 10 to Figure 13. The most significant work in the Peace on agricultural settlement has been carried out by Tracie (1970). He suggests that vegetation cover appears to be the most probable determinant of farm location in the period 1908 to 1927.

Apparently, homesteaders were most attracted by the lighter aspen poplar parkland areas, where there was at least 20% grassland. From 1928 to 1968, the proximity of the homestead to settlement centres and major transportation links becomes more important.

The steady gain in population of the study area seems to have been caused by the natural excess of births over deaths, with the exception of the period 1921 to 1931, when there was a net migration gain. In the other three decades, estimates indicate a net migration loss occurred, in which out-migration exceeded in-migration (Table 4.2).

TABLE 4.2 - POPULATION CHANGE BY TIME PERIOD FOR STUDY AREA

	Population Increase		Natural Increase		Migration Gain or Loss	
	Absolute	%*	Absolute	%*	Absolute	%*
1921-1931	21,022	130.98	3,757	23.41	+17,265	+107.57
1931-1941	5,467	14.75	7,486	20.19	- 2,019	- 5.45
1941-1951	5,939	13.96	10,564	24.83	- 4,625	- 10.87
1951-1961	10,776	22.23	17,941	37.01	- 7,165	- 14.78

*The percentage figure is obtained by expressing the absolute gain or loss for the time period as a percentage of the population at the beginning of the time period.

Most regions of the study area have experienced steady growth over the period 1921 to 1961 (Figures 14, 16, 18, 20). The exceptions appear to be the less densely settled areas on the margins of the study area. Appendix C provides a full description of the population changes in each of the twenty-six regions. The location of these regions are

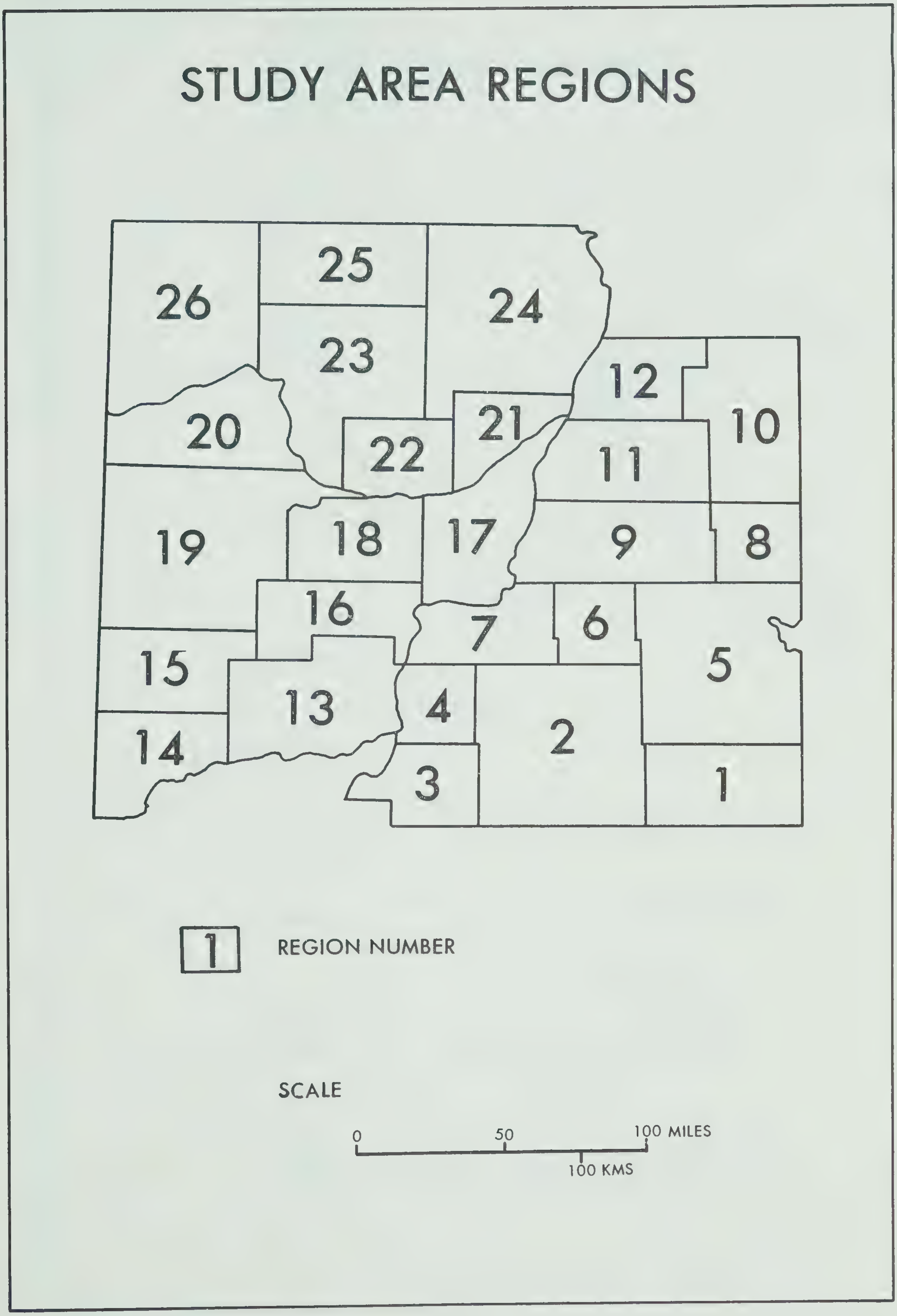

STUDY AREA REGIONS
26
25
24
23
20
22
21
12
10
11
19
18
17
9
8
16
7
6
5
15
13
4
2
14
3
1
1
REGION NUMBER
SCALE
0
50
100 MILES
100 KMS

FIGURE 9

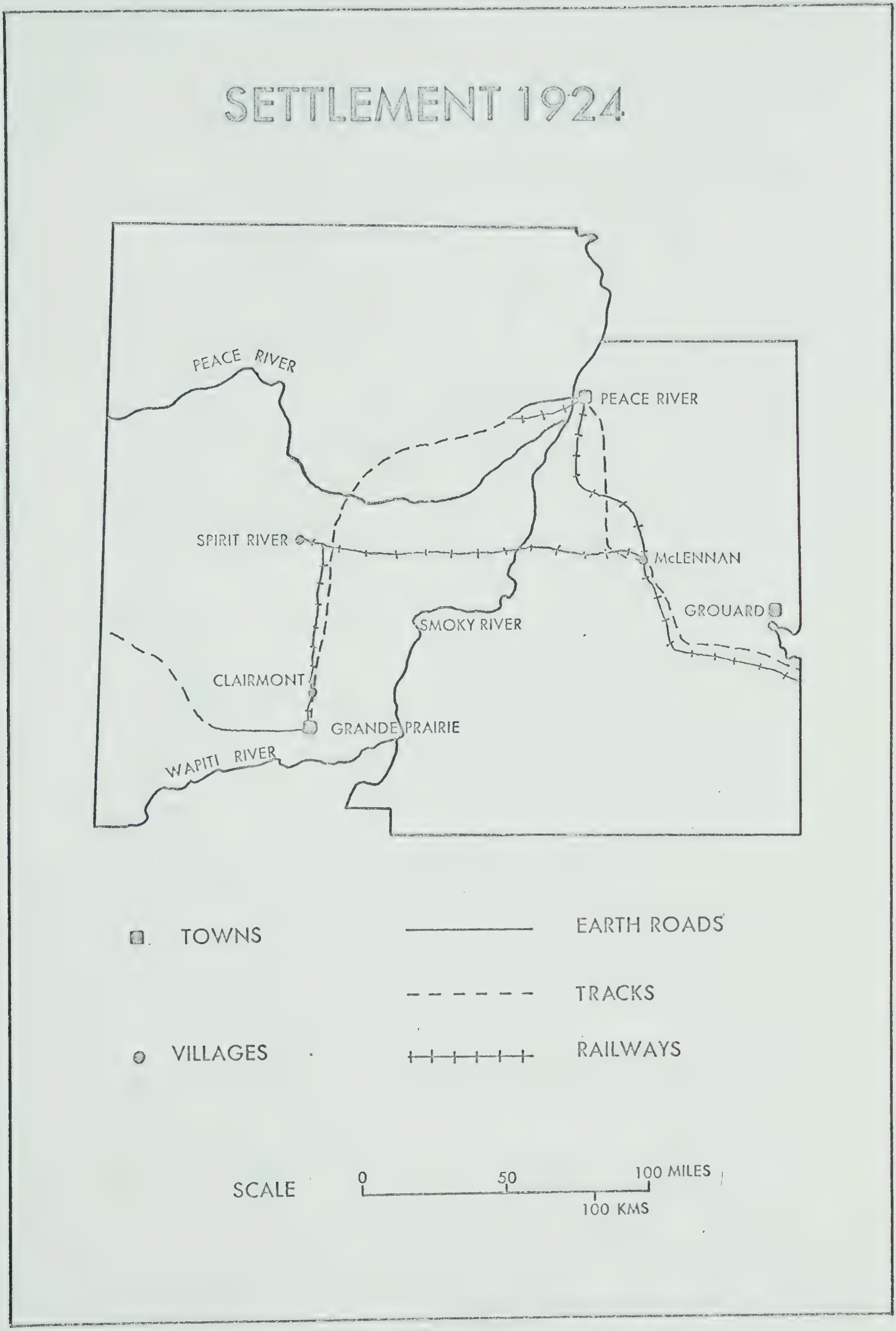

SETTLEMENT 1924
PEACE RIVER
PEACE RIVER
SPIRIT RIVER
McLENNAN
GROUARD
SMOKY RIVER
CLAIRMONT
GRANDE PRAIRIE
WAPITI RIVER
TOWNS
EARTH ROADS
TRACKS
VILLAGES
RAILWAYS
SCALE
0
50
100 MILES
100 KMS

FIGURE 10

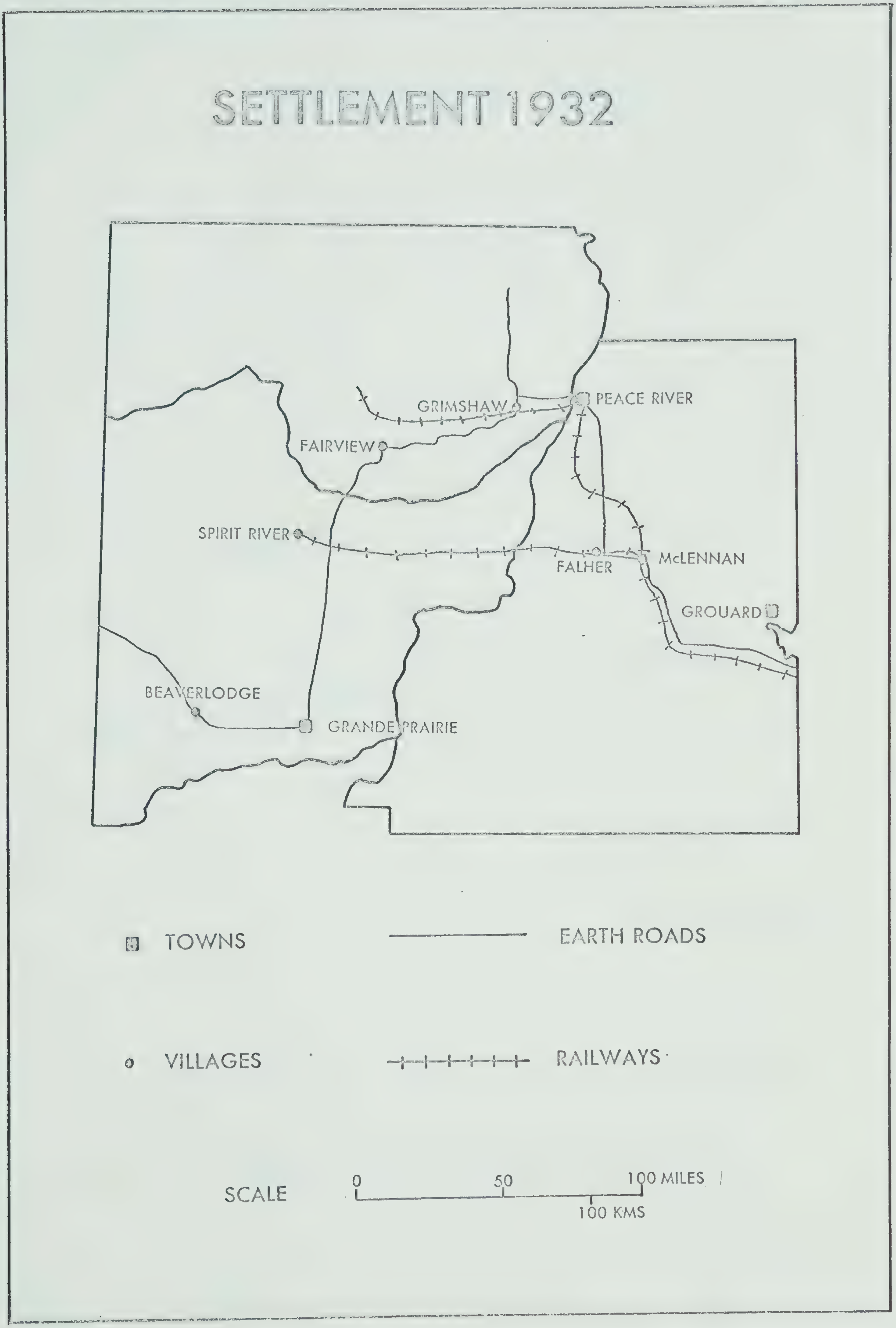

SETTLEMENT 1932
PEACE RIVER
GRIMSHAW
FAIRVIEW
SPIRIT RIVER
FALHER
McLENNAN
GROUARD
BEAVERLODGE
GRANDE PRAIRIE
TOWNS
EARTH ROADS
VILLAGES
RAILWAYS
SCALE
0
50
100 MILES
100 KMS

FIGURE 11

SETTLEMENT 1948

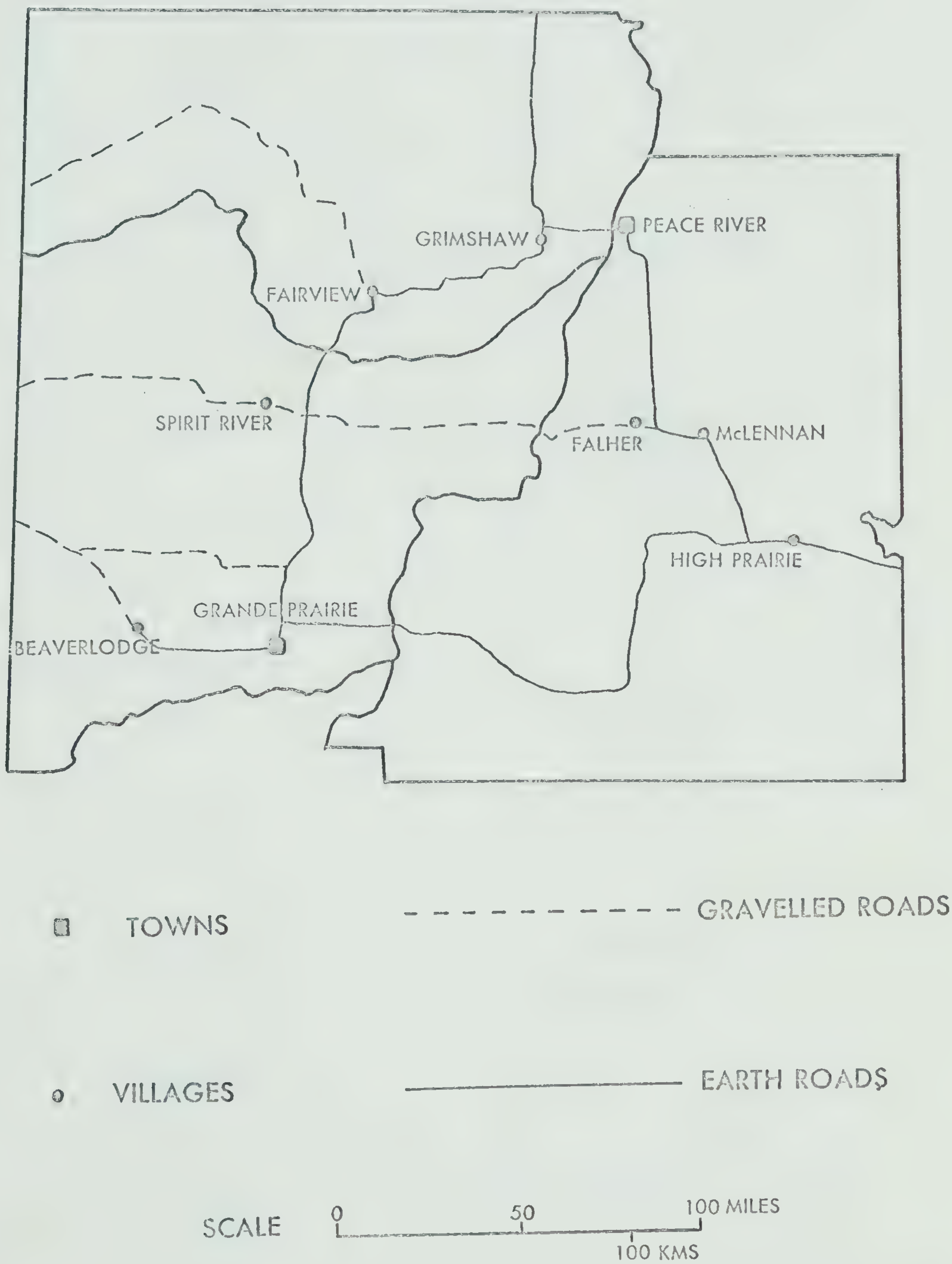


FIGURE 12

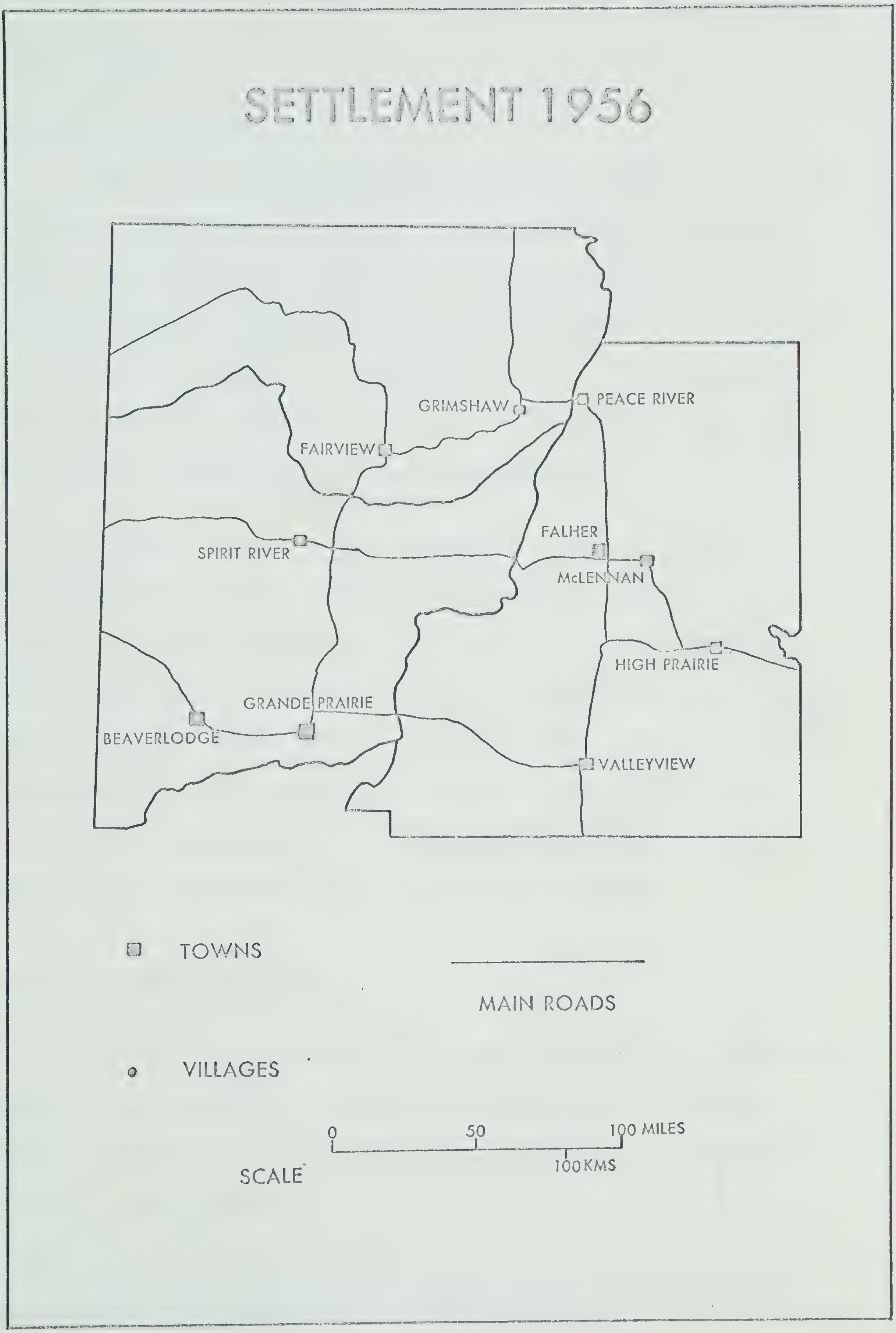

SETTLEMENT 1956
GRIMSHAW
PEACE RIVER
FAIRVIEW
FALHER
SPIRIT RIVER
McLENNAN
HIGH PRAIRIE
GRANDE PRAIRIE
BEAVERLODGE
VALLEYVIEW
TOWNS
MAIN ROADS
VILLAGES
0
50
100 MILES
100 KMS
SCALE

FIGURE 13

given in Figure 9.

3. Relevant Literature

There is a variety of standard historical works on the history of the Peace River country (MacGregor 1952, Kitto 1930, Wellman 1965). The most significant work in respect to this study has been carried out by geographers. Tracie (1967) defined some of the factors involved in the migration of people into the Southern Peace[3]. These included favourable land costs, physical advantages, social contacts, improved communications, and Government Homesteading policies.

Marriott's (1969) study of migration within the County of Grande Prairie for the years 1956 to 1967 attempted to suggest some of the significant factors determining local migration. Multivariate analysis in his study was unable to suggest any significant relationship between migration distance and independent variables, other than the size of the population centre. Such a relationship between distance and population size is summed up by both the gravity model and the Pareto functions used in the Markov chain and Monte Carlo models.

Inputs to the Monte Carlo Model

The basic requirements of the Monte Carlo model are population and area characteristics of the twenty-six regions, the number of potential migrants from each region, and a statement of the expected interaction between regions.

---

[3]The Southern Peace is that district bounded by the British Columbian border, the Wapiti, Smoky and Peace Rivers.

Area characteristics are provided in Table 4.1. Distance between regions is calculated by a routine which measures the interval between the centres of regions, where the location of each centre is expressed in geographic co-ordinates. Adjusted population, Ap values (Table 4.3), are estimated by expression (2.8). Natural increase and external migration figures required in the calculation of the adjusted populations of regions are presented in Appendix C.

The number of potential migrants, mj , from each region is estimated as a percentage of the adjusted population (Table 4.3). The percentage values for each time period are calculated empirically from sample data. A similar procedure was followed by Morrill in Sweden (1965b, p. 155). In time period one, 1921-1931, a value of 15.40% is used. This implies that during the decade 15.40% of the adjusted population for each region are expected to move internally. Figures for the periods, 1931-1941, 1941-1951, and 1951-1961 are respectively 19.50%, 25.90%, and 24.00%. The amount of local migration generally increases over time. Morrill's values behave similarily.

The basic interaction between regions is determined by means of Pareto functions, calculated by the method presented in Appendix A. These functions (Table 4.4) are utilised in the model through expression (2.6). Exponent values have frequently been used to represent graphically the extent of the Mean Information Field. Figure 4 indicates that the migration fields vary over time, but not widely. The exponents are larger than the b value of 1 , which is often proposed for North American studies. They are closer to the results of Swedish studies, which indicated exponents varying from -2 to -2.5 for rural

TABLE 4.3 - ADJUSTED POPULATIONS AND POTENTIAL MIGRANTS OF STUDY REGIONS*

Region	1921-1931		1931-1941	
	Adjusted Population	Potential Migrants	Adjusted Population	Potential Migrants
1	24	4	28	5
2	556	86	620	121
3	26	4	107	21
4	420	65	555	108
5	2817	434	3128	610
6	617	95	590	115
7	24	4	134	26
8	155	24	157	31
9	3524	543	4025	785
10	10	1	0	0
11	1328	205	1230	240
12	1395	215	1241	242
13	6975	1075	7931	1548
14	2298	354	1819	355
15	3254	501	3609	704
16	1407	217	2095	409
17	722	111	1084	211
18	2475	381	2569	501
19	1379	212	2056	401
20	33	5	41	8
21	2192	338	2319	452
22	2063	318	2748	536
23	1764	272	2150	419
24	1397	215	1813	354
25	24	4	92	18
26	193	30	398	78
Total	37072	5713	42539	8298

*See Chapter Two, Section IV, for the definitions of Adjusted Population and Potential Migrants.

TABLE 4.3 (Continued)

Region	1941-1951 Adjusted Population	1941-1951 Potential Migrants	1951-1961 Adjusted Population	1951-1961 Potential Migrants
1	24	6	114	27
2	1114	289	2267	543
3	31	8	63	15
4	816	211	1165	279
5	4200	1088	4213	1009
6	964	250	932	223
7	66	17	127	30
8	176	46	367	88
9	4970	1288	5516	1322
10	13	3	1	0
11	1649	427	1915	459
12	1839	476	3595	861
13	7927	2054	3188	3161
14	1108	287	1588	380
15	3965	1027	4506	1080
16	1355	351	169	384
17	2356	610	1820	436
18	3007	779	3986	955
19	1999	518	1257	301
20	127	33	287	69
21	2711	702	3269	783
22	2964	768	3789	908
23	2365	613	1304	312
24	2276	590	1724	413
25	98	25	91	22
26	358	93	569	136
Total	48478	12559	59254	14196

TABLE 4.4 - EXPONENT AND INTERCEPT VALUES OF PARETO FUNCTIONS USED IN THE MONTE CARLO SIMULATION

	Exponent	Intercept
1921-1931	- 2.80	5.370
1931-1941	- 2.46	1.047
1941-1951	- 2.77	4.787
1951-1961	- 2.18	0.692

populations (Hagerstrand 1957, p. 118). Olsson (1965b, p. 60) suggested that most exponent values fall between -1.74 and -3.00, for all types of interaction.

Due to the lack of relevant information pertaining to the influence of social and economic factors through time, detailed interpretation of these exponent values is not possible here. Time period one (1921-1931) has the steepest migration field, -2.80, though there is little difference between it and time three (1941-1951), -2.77. In contrast, periods two (1931-1941) and four (1951-1961) have slightly more extensive fields.

In practice, the symmetrical inverse-distance relationship is modified by use of absolute barriers to movement (Table 4.5). These barriers were constructed empirically from sample evidence. They reflect the spread of settlement through time, and the presence of unfavourable physical conditions. For example, in time period one, 1921-1931, no region is permitted to send migrants to region 26, since actual settlement during that decade was sparse. After 1931 settlement in this area expanded, and the barriers are accordingly removed. While the barriers

TABLE 4.5 - EFFECTIVE BARRIERS TO MIGRATION BETWEEN REGIONS

Region to which barrier is effective	Regions from which barriers are effective			
	1921-1931	1931-1941	1941-1951	1951-1961
1	All regions	All regions except 2 and 5	All regions except 2 and 5	
3	All regions	All regions except 2 and 4	All regions except 2 and 4	All regions except 2 and 4
7	All regions	All regions except 2,4, 6 and 9	All regions except 2,4, 6 and 9	
10	All regions	All regions except 11	All regions	All regions except 11
19	All regions except 15, 16 and 18			
20	All regions	All regions except 19, 23 and 26	All regions except 19, 23 and 26	
25	All regions	All regions except 23, 24 and 26	All regions except 23, 24 and 26	All regions except 23, 24 and 26
26	All regions			

have little theoretical meaning, they impart a degree of flexibility to the Monte Carlo probabilities.

Monte Carlo Simulation

The raw probability of contact, estimated by the Pareto expression, is converted to smoothed and adjusted values according to expression (2.9). An example of derived probabilities for one region is given in Table 4.6. The Monte Carlo programme computes internally a range of random numbers relative to the adjusted probability value. A random number is drawn for each potential migrant in the region. The location of each migrant is memorised. In this particular example, no barriers were encountered, so no random numbers have to be regenerated. Ideally, the sequence of random numbers, $\{R_n\}$, should have a serial correlation equal to zero, reflecting the degree of dependence between succeeding terms. The routine used to draw numbers has a serial correlation of the order 1.5×10^{-5}, and is considered adequate for this study.

Similar tables to 4.6 are produced for all regions in the simulation run. The programme presents the gain or loss of migrants for all regions in the form of a 26×26 chart. Table 4.7 is the result of one complete simulation run for the period 1921 to 1931. Clustering of the largest values about the diagonal indicates the greater importance of short distance moves, many of which do not cross the borders of the region of origin. Regions with the largest number of values in the columns receive the heaviest migration gains. Region 13, which contains Grande Prairie, the largest population centre, clearly has a marked migration gain, as does region 18 in which Spirit River, an early growth point for the area, is located.

TABLE 4.6 - ABBREVIATED EXAMPLE OF MONTE CARLO PROBABILITIES, REGION ONE IN TIME PERIOD 1921-1931

Region of Origin	Region of Destination	Raw Probability	Smoothed Probability	Adjusted Probability	Population of Desintation Region	Distance Between Regions (Kilometres)
1	1	1.37	0.6125	0.0183	24	23.051*
1	2	0.18	0.0786	0.0543	556	61.474
1	3	0.01	0.0052	0.0002	26	106.702
1	4	0.01	0.0038	0.0020	420	107.344
1	5	0.48	0.2154	0.7542	2817	42.689
1	6	0.03	0.0120	0.0092	617	73.394
1	7	0.01	0.0067	0.0002	24	101.462
1	8	0.02	0.0072	0.0014	155	88.127
1	9	0.03	0.0136	0.0594	3524	95.144
1	10	0.01	0.0051	0.0001	10	131.952
1	11	0.01	0.0059	0.0098	1328	121.660
1	12	0.00	0.0021	0.0037	1395	149.719
1	13	0.01	0.0044	0.0384	6975	148.687
1	14	0.00	0.0011	0.0032	2298	205.885
1	15	0.00	0.0012	0.0050	3254	203.478
1	16	0.01	0.0025	0.0043	1407	153.663
1	17	0.01	0.0049	0.0044	722	126.902

*The distance between region 1 and region 1 is the radius of the circle analogous in area to region 1.

TABLE 4.6 (Continued)

Region of Origin	Region of Destination	Raw Probability	Smoothed Probability	Adjusted Probability	Population of Destination Region	Distance Between Regions (Kilometres)
1	18	0.01	0.0023	0.0070	2475	158.263
1	19	0.01	0.0030	0.0051	1379	208.483
1	20	0.00	0.0011	0.0000	33	224.895
1	21	0.00	0.0017	0.0046	2192	152.056
1	22	0.00	0.0015	0.0037	2063	166.052
1	23	0.00	0.0021	0.0047	1764	200.243
1	24	0.01	0.0038	0.0066	1397	185.343
1	25	0.00	0.0010	0.0000	24	229.086
1	26	0.00	0.0014	0.0003	193	258.104

Total Migrants = 4

Regenerated 0 migration[+]

Denominator for adjusted population[++] 804.67

[+] No random numbers had to be regenerated, since no barriers were encountered.

[++] The denominator is equivalent to $P_i \cdot H_i / \sum P_i \cdot H_i$.

TABLE 4.7 - EXAMPLE OF A SINGLE SIMULATION ORIGIN BY DESTINATION CHART, TIME PERIOD 1921-1931

Region of Origin	Region of Destination																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1		1			3																					
2		30			25	4			10		2		9	1		4		1								
3		1											3													
4		4		9		2			6			1	37		1	2	3									
5		5			337	9		1	57		8	3	6				5	2			1					
6		2			21	21			43		2		4				2									
7		1							1				2													
8					11			3	10																	
9		3			37	16		1	354		80	10	9		3	1	9	5			9	4	2			
10					1																					
11		2			3				98		58	22	3	1	2		3	4			6		1	2		
12		2			1				22		44	108	1		1		1				11	2	5	17		
13		4		5	5	1			8				881	23	44	78	2	21			1		2			
14		2			1								43	208	97	2	1									
15		1		1	1								67	77	301	13	1	7	32							
16		2							1				115	1	6	41	1	41	6		2	1				
17		2		1	1	5			21		1	3	10		1	4	29	15			6	10	2			
18		3							3		3		47	1	6	65	4	179	15		6	40	9			
19		2							1				34	8	36	13	2	15	92			4	5			
20													1					2		1			1			
21		3			2				14		16	6	6	2	2		17	8			182	21	20	39		
22		2		1					5			1	12	3	1	7	8	58			13	148	53	6		
23		2		1					2		2	1	11		4	3	6	12			11	42	157	18		
24		2			1				10		9	9	5		1	1	4	5			27	7	21	113		
25												1											3			
26											2	2	2	1			1	1			2		15			4

The column totals provide the net migration gain, the row totals, the loss. Results of four single simulations through time are given in Table 4.8. Each single simulation was recorded, so that the average net gain or loss by internal migration could be calculated. In total, fifty simulations in each time period were averaged out. The means and standard deviations of the averaged distributions are listed in Table 4.9.

The final step involves the specification of the new population distribution, using expression (2.13). The internal migration gain or loss, gk, for each region, is added to the adjusted population, Ap , to give a new distribution, Np (Table 4.10). Averaged simulated population distributions are presented in cartographic form in Figures 15, 17, 19 and 21, and can be visually compared with the actual population changes (Figures 14, 16, 18 and 20). Simulated results are tested against actual observations in Chapter Five, as are the equilibrum solutions achieved by the Markov chain model.

Markov Chain Inputs and Outputs

Equilibrium population distributions from the Markov chain model were determined by the procedure illustrated in Chapter Three. Transition matrices (Tables 4.12 to 4.15), indicating the probability of movement between regions i and j , are constructed by means of a gravity model (Expression 3.19). The illustrated matrices show only n - 1 states. The nth state, which has been presented separately (Table 4.15), consists of the region's death rate (D value). In some cases the D value is taken as the region's net population decrease, to

TABLE 4.8 - GAIN OR LOSS OF MIGRANTS BY REGION: SINGLE SIMULATION RUN

Region	1921-1931			1931-1941			1941-1951			1951-1961		
	Gain	Loss	Total	Gain	Loss	Total	Gain	Loss	Total	Gain	Loss	Total
1	0	4	- 4	0	5	- 5	1	6	- 5	20	27	- 7
2	76	86	-10	74	121	-47	237	289	-52	581	543	38
3	0	4	- 4	4	21	-17	1	8	- 7	2	15	-13
4	18	65	-47	39	108	-69	97	211	-114	104	279	-175
5	450	434	16	636	610	26	1218	1088	130	1130	1009	121
6	58	95	-37	63	115	-52	152	250	-98	101	223	-122
7	0	4	- 4	2	26	-24	2	17	-15	12	30	-18
8	5	24	-19	3	31	-28	7	46	-39	24	88	-64
9	666	543	123	916	785	131	1586	1288	298	1557	1322	235
10	0	1	- 1	0	0	0	0	3	- 3	0	0	0
11	227	205	22	252	240	12	439	427	12	514	459	55
12	167	215	-48	146	242	-96	368	476	-108	715	861	-146
13	1308	1075	233	1941	1548	393	2530	2054	476	4028	3161	867
14	326	354	-28	272	355	-83	184	287	-103	224	380	-156
15	506	501	5	688	704	-16	987	1027	-40	942	1080	-138
16	234	217	17	367	409	-42	297	351	-54	327	384	-57
17	99	111	-12	177	211	-34	610	610	0	390	436	-46
18	376	381	- 5	469	501	-32	776	779	- 3	890	955	-65
19	145	212	-67	496	401	95	547	518	29	310	301	9
20	1	5	- 4	1	8	- 7	5	33	-28	10	69	-50
21	277	338	-61	341	452	-111	559	702	-143	613	783	-170
22	279	318	-39	477	536	-59	681	768	-87	797	908	-111
23	296	272	24	473	419	54	620	613	7	283	312	-29
24	195	215	-20	396	354	42	613	590	23	514	413	101
25	0	4	- 4	7	18	-11	5	25	-20	1	22	-21
26	4	30	-26	58	78	-20	37	93	-56	98	136	-38
Total	5713	5713	0	8298	8298	0	12,559	12,559	0	14,196	14,196	0

TABLE 4.9 - MEAN AND STANDARD DEVIATION OF GAIN OR LOSS OF MIGRANTS BY REGION: FIFTY SIMULATIONS

Region	1921-1931		1931-1941		1941-1951		1951-1961	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
1	- 3.520	0.839	- 2.820	2.201	- 4.320	1.504	- 15.320	7.325
2	-16.860	19.339	- 10.860	32.450	- 38.540	29.176	19.520	43.760
3	- 3.980	0.141	- 18.660	1.409	- 6.800	1.069	- 13.240	1.255
4	-40.420	6.078	- 66.420	10.051	-114.100	10.895	-159.020	15.802
5	46.080	18.193	58.720	27.761	122.240	25.878	120.380	31.707
6	-42.960	7.088	- 60.780	7.765	-112.740	10.369	-109.780	10.739
7	- 3.900	0.303	- 21.240	1.836	- 14.220	1.569	- 15.580	3.631
8	-18.260	1.998	- 23.500	2.765	- 35.680	3.087	- 60.560	5.128
9	110.000	17.580	134.200	23.961	268.080	35.254	251.280	28.016
10	- 1.000	0.000	0.000	0.000	- 3.000	0.000	0.020	0.141
11	11.380	12.602	- 2.360	12.184	8.540	17.193	34.900	19.114
12	-52.540	8.474	- 90.680	9.757	-132.540	11.511	-157.580	15.350
13	266.700	33.096	400.880	49.097	444.680	43.181	873.160	53.644
14	-50.200	16.999	-101.020	14.178	-102.500	12.397	-139.640	14.067
15	- 6.420	17.278	- 50.260	20.069	- 45.020	18.230	-125.080	20.491
16	-15.500	11.689	- 18.100	15.635	- 33.460	18.686	- 39.440	16.335
17	-23.520	10.516	- 30.960	12.415	- 20.380	20.526	- 42.860	16.862
18	-17.080	13.930	- 39.460	16.028	- 36.960	16.223	- 68.480	20.660
19	-60.200	10.660	68.560	21.684	40.200	21.718	11.060	17.303
20	- 4.820	0.388	- 5.800	1.325	- 22.900	3.005	- 36.900	5.807
21	-48.320	16.154	-106.660	15.141	-122.760	21.731	-156.560	22.710
22	-32.820	11.832	- 72.020	18.859	- 73.360	22.940	-117.080	22.410
23	16.520	20.367	37.320	20.822	27.780	24.216	- 23.700	19.864
24	18.080	26.708	69.440	34.934	75.680	43.711	46.560	40.489
25	- 3.980	0.141	- 12.640	2.345	- 18.280	2.466	- 17.500	2.358
26	-22.460	2.252	- 34.880	13.563	- 49.640	8.935	- 58.560	18.970

TABLE 4.10 - RESULTS OF RANDOMLY CHOSEN SINGLE SIMULATION RUNS BY TIME PERIOD

Region	1931		1941		1951		1961	
	Absolute	Percent	Absolute	Percent	Absolute	Percent	Absolute	Percent
1	20	0.054	23	0.054	19	0.039	107	0.181
2	546	1.473	573	1.347	1062	2.191	2305	3.890
3	22	0.059	90	0.212	24	0.049	50	0.084
4	373	1.006	486	1.142	702	1.448	990	1.671
5	2833	7.642	3154	7.414	4330	8.932	4334	7.314
6	580	1.564	538	1.265	866	1.786	810	1.367
7	20	0.054	110	0.265	51	0.105	109	0.184
8	136	0.367	129	0.303	137	0.283	303	0.511
9	3647	9.838	4156	9.770	5268	10.867	5751	9.706
10	9	0.024	0	0.000	10	0.021	1	0.002
11	1350	3.642	1242	2.920	1661	3.426	1970	3.325
12	1347	3.633	1145	2.692	1731	3.571	3449	5.821
13	7208	19.443	8324	19.568	8403	17.334	14055	23.720
14	2270	6.123	1736	4.081	1005	2.073	1432	2.417
15	3259	8.791	3593	8.446	3925	8.096	4368	7.372
16	1424	3.841	2053	4.826	1301	2.684	1544	2.606
17	710	1.915	1050	2.468	2356	4.860	1774	2.994
18	2470	6.663	2537	5.964	3004	6.197	3921	6.617
19	1312	3.539	2151	5.056	2028	4.183	1266	2.137
20	29	0.078	34	0.080	99	0.204	237	0.400
21	2131	5.748	2208	5.190	2568	5.297	3099	5.230
22	2024	5.460	2689	6.321	2877	5.935	3678	6.207
23	1788	4.823	2204	5.181	2372	4.893	1275	2.152
24	1377	3.714	1855	4.361	2299	4.742	1825	3.080
25	20	0.054	81	0.190	78	0.161	70	0.118
26	167	0.450	378	0.889	302	0.623	531	0.896
Total	37,072	100.000	42,539	100.000	48,478	100.000	59,254	100.000

Calculation by the author.

TABLE 4.11 - POPULATION OF STUDY REGIONS PREDICTED BY AVERAGED SIMULATION SOLUTION

Region	1931		1941		1951		1961	
	Absolute	Percent	Absolute	Percent	Absolute	Percent	Absolute	Percent
1	21	0.057	25	0.059	20	0.041	99	0.167
2	539	1.454	609	1.432	1076	2.220	2286	3.858
3	22	0.059	88	0.207	24	0.049	50	0.084
4	380	1.025	489	1.149	702	1.448	1006	1.698
5	2863	7.723	3187	7.492	4322	8.915	4333	7.313
6	574	1.548	529	1.244	851	1.755	822	1.387
7	20	0.054	113	0.266	52	0.107	111	0.187
8	137	0.369	134	0.315	140	0.289	306	0.516
9	3633	9.800	4159	9.777	5238	10.805	5767	9.733
10	9	0.024	0	0.000	10	0.021	1	0.002
11	1339	3.612	1228	2.887	1657	3.418	1950	3.291
12	1343	3.623	1150	2.703	1707	3.521	3437	5.800
13	7240	19.530	8332	19.587	8370	17.266	14063	23.733
14	2248	6.064	1718	4.039	1006	2.075	1448	2.444
15	3248	8.761	3559	8.366	3920	8.086	4381	7.394
16	1392	3.755	2077	4.883	1322	2.727	1562	2.636
17	699	1.885	1053	2.475	2336	4.819	1777	2.999
18	2458	6.630	2530	5.947	2970	6.126	3918	6.612
19	1319	3.558	2125	4.995	2039	4.206	1268	2.140
20	28	0.075	35	0.082	104	0.214	250	0.422
21	2144	5.783	2212	5.200	2588	5.338	3112	5.252
22	2030	5.476	2676	6.291	2891	5.963	3672	6.197
23	1780	4.801	2187	5.141	2393	4.936	1280	2.160
24	1415	3.817	1882	4.424	2352	4.852	1771	2.989
25	20	0.054	79	0.186	80	0.165	74	0.125
26	171	0.461	363	0.853	308	0.635	510	0.861
Total	37,072	100.000	42,539	100.000	48,478	100.000	59,254	100.000

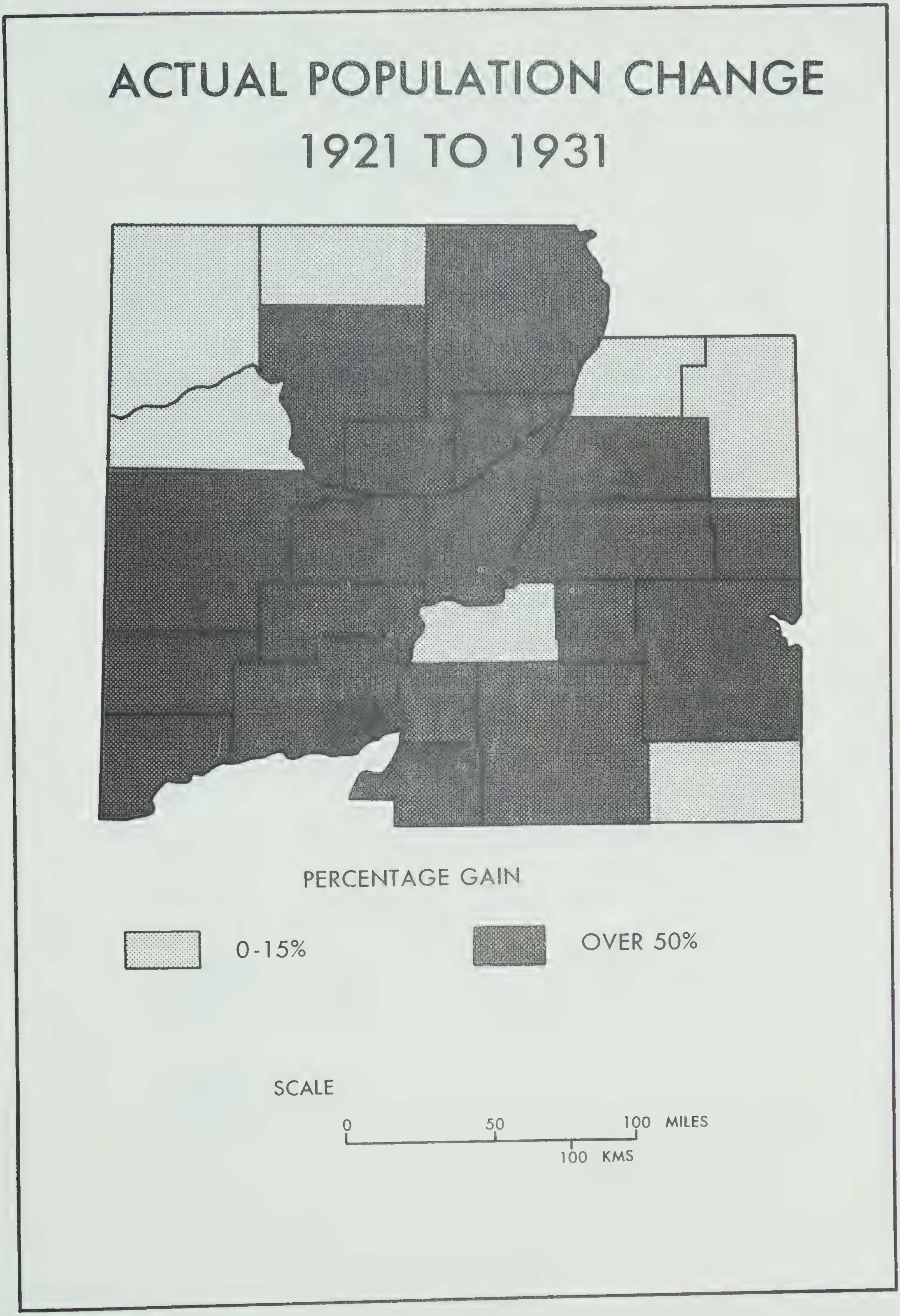

ACTUAL POPULATION CHANGE
1921 TO 1931
PERCENTAGE GAIN
0-15%
OVER 50%
SCALE
0
50
100 MILES
100 KMS

FIGURE 14

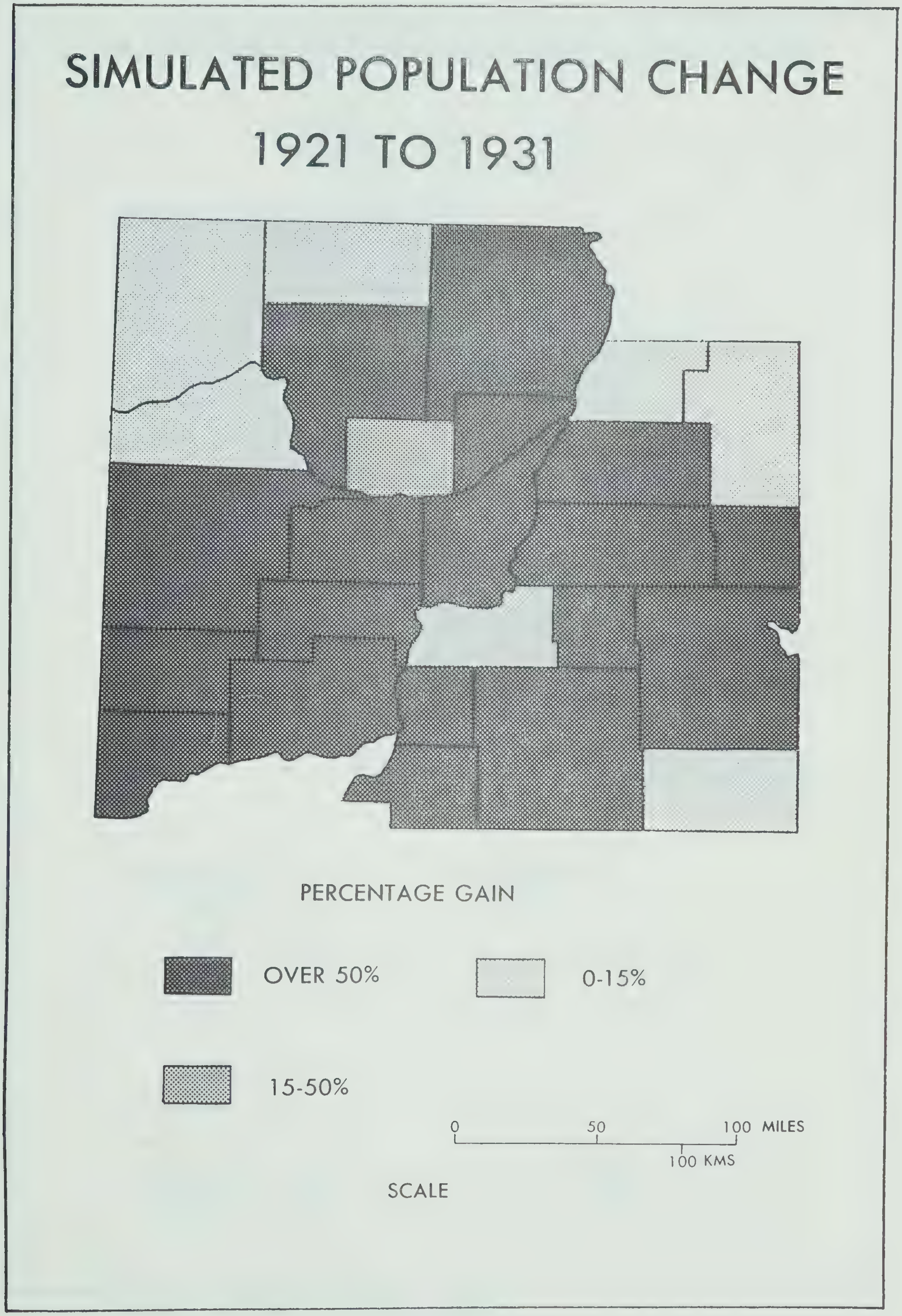

SIMULATED POPULATION CHANGE
1921 TO 1931
PERCENTAGE GAIN
OVER 50%
0-15%
15-50%
0
50
100 MILES
100 KMS
SCALE

FIGURE 15

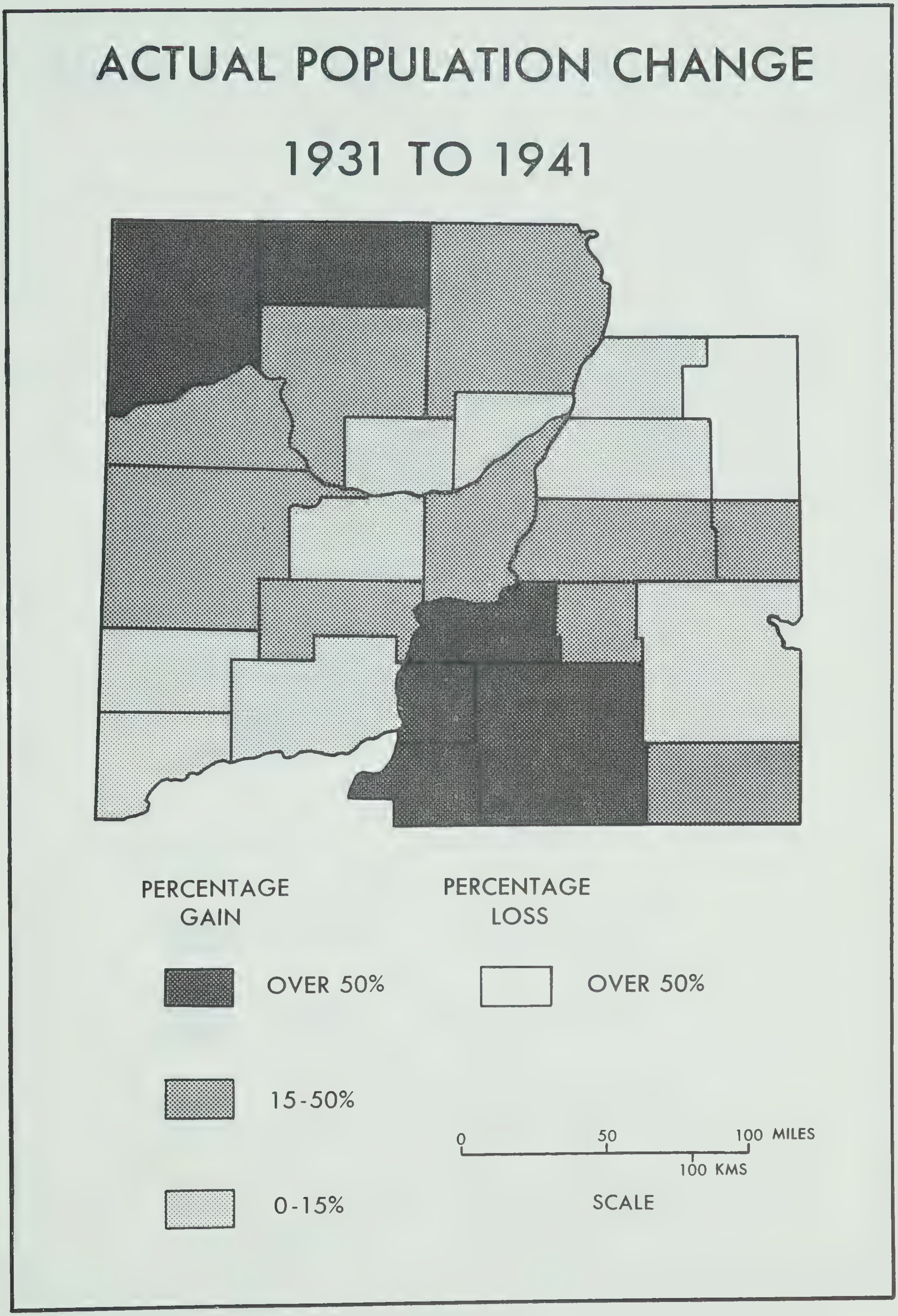

ACTUAL POPULATION CHANGE
1931 TO 1941
PERCENTAGE GAIN
OVER 50%
15-50%
0-15%
PERCENTAGE LOSS
OVER 50%
0
50
100 MILES
100 KMS
SCALE

FIGURE 16

SIMULATED POPULATION CHANGE 1931 TO 1941

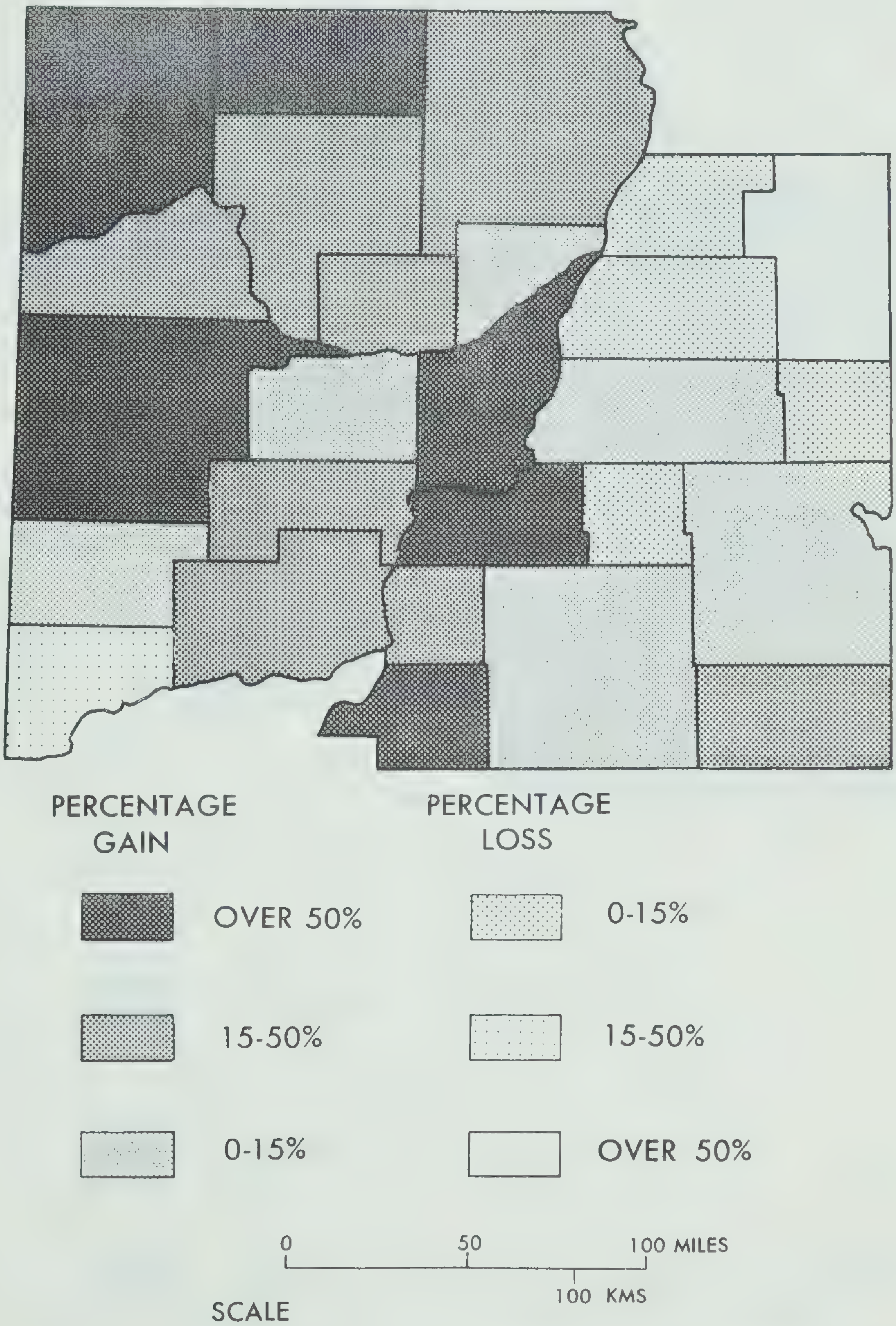


FIGURE 17

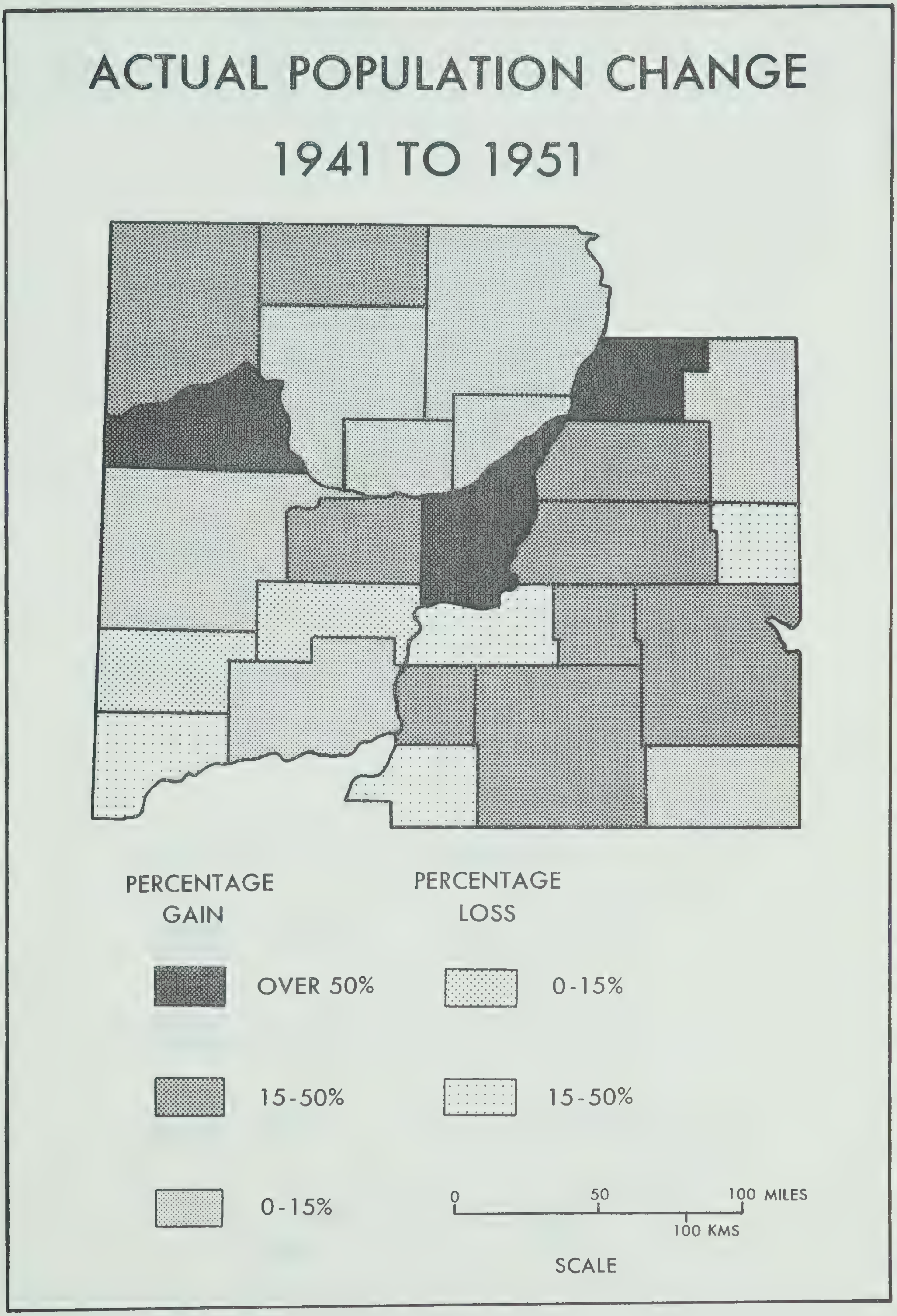

ACTUAL POPULATION CHANGE
1941 TO 1951
PERCENTAGE GAIN
OVER 50%
15-50%
0-15%
PERCENTAGE LOSS
0-15%
15-50%
0
50
100 MILES
100 KMS
SCALE

FIGURE 18

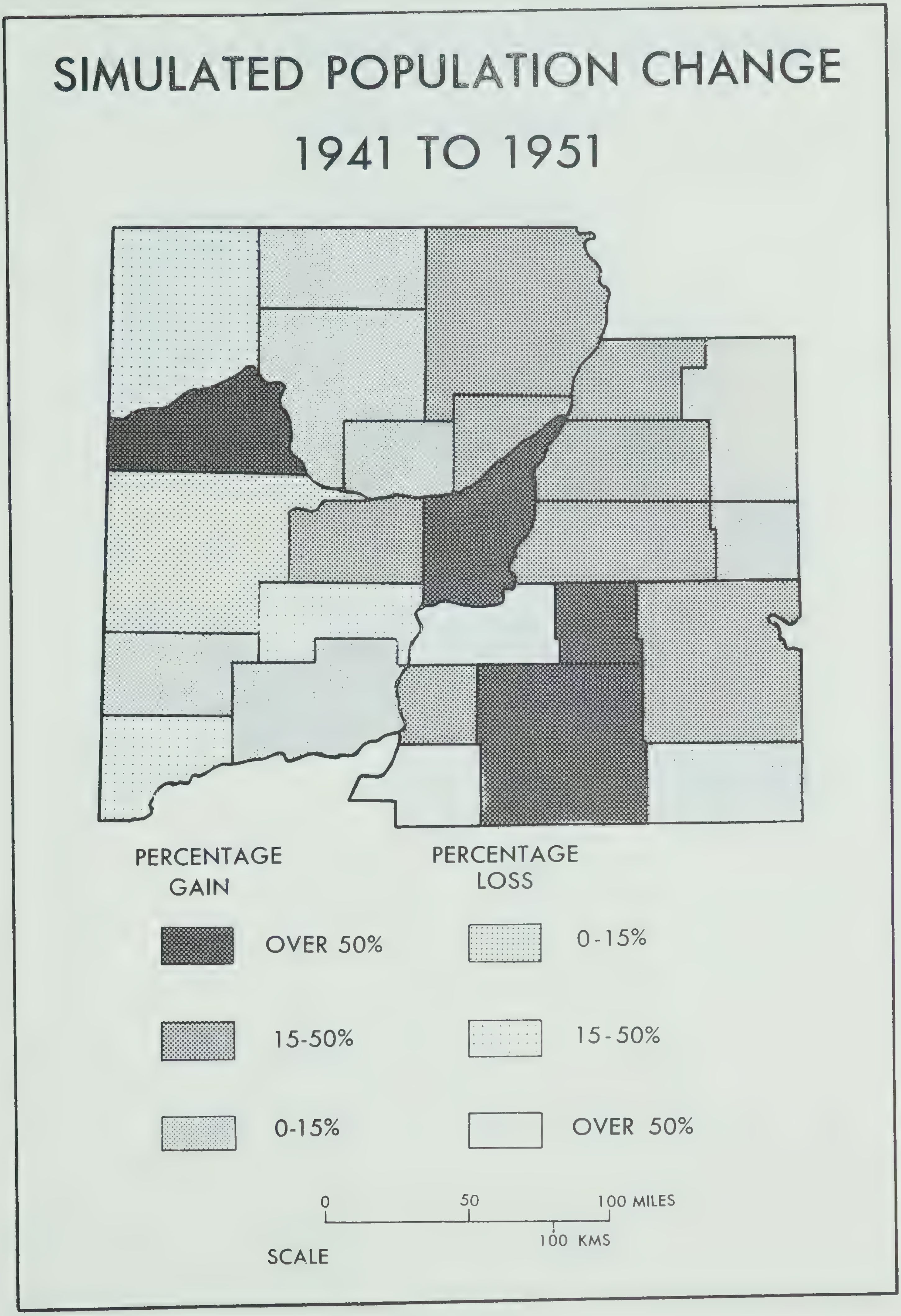

SIMULATED POPULATION CHANGE
1941 TO 1951
PERCENTAGE GAIN
OVER 50%
15-50%
0-15%
PERCENTAGE LOSS
0-15%
15-50%
OVER 50%
0
50
100 MILES
100 KMS
SCALE

FIGURE 19

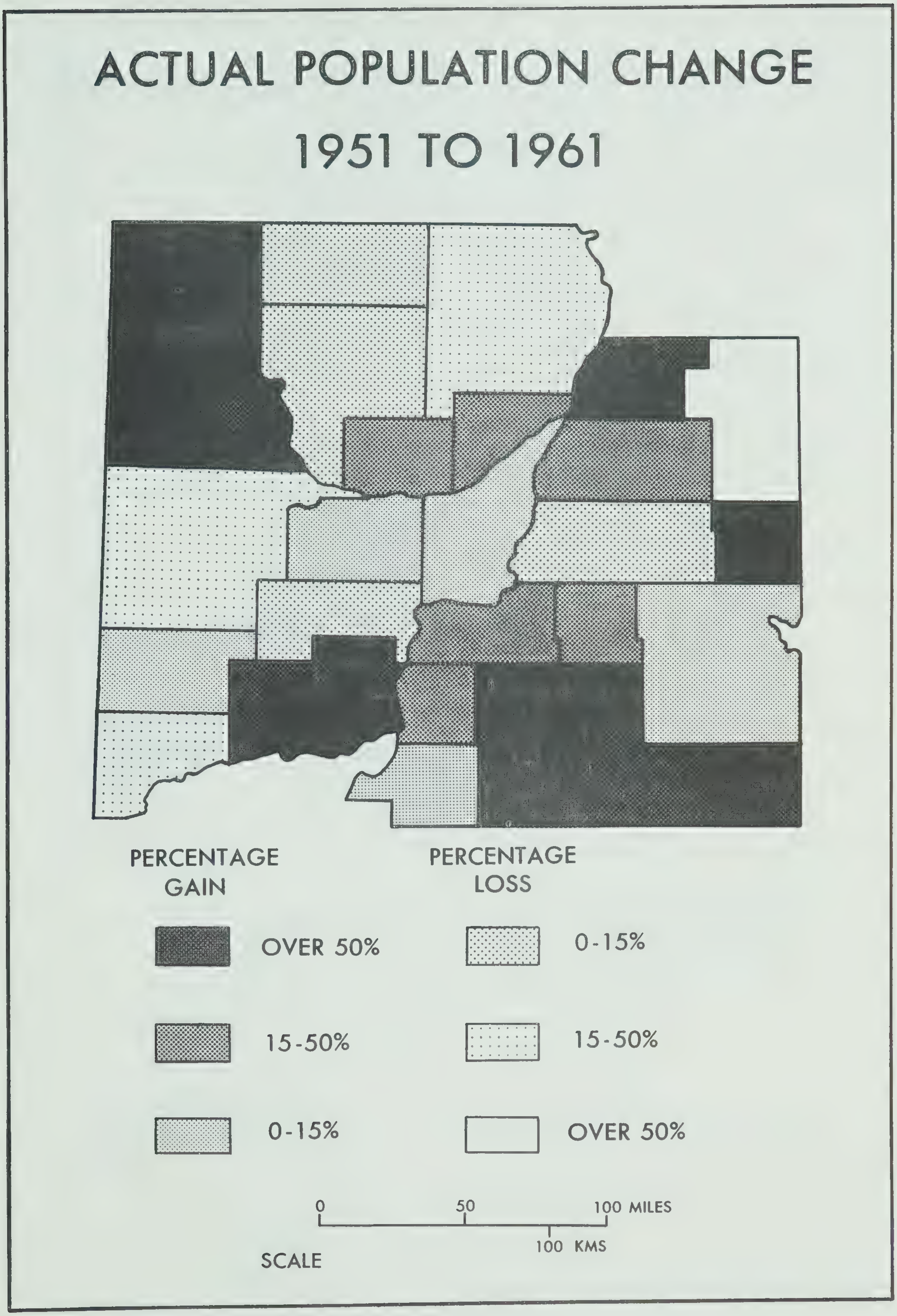

ACTUAL POPULATION CHANGE
1951 TO 1961
PERCENTAGE GAIN
OVER 50%
15-50%
0-15%
PERCENTAGE LOSS
0-15%
15-50%
OVER 50%
0
50
100 MILES
100 KMS
SCALE

FIGURE 20

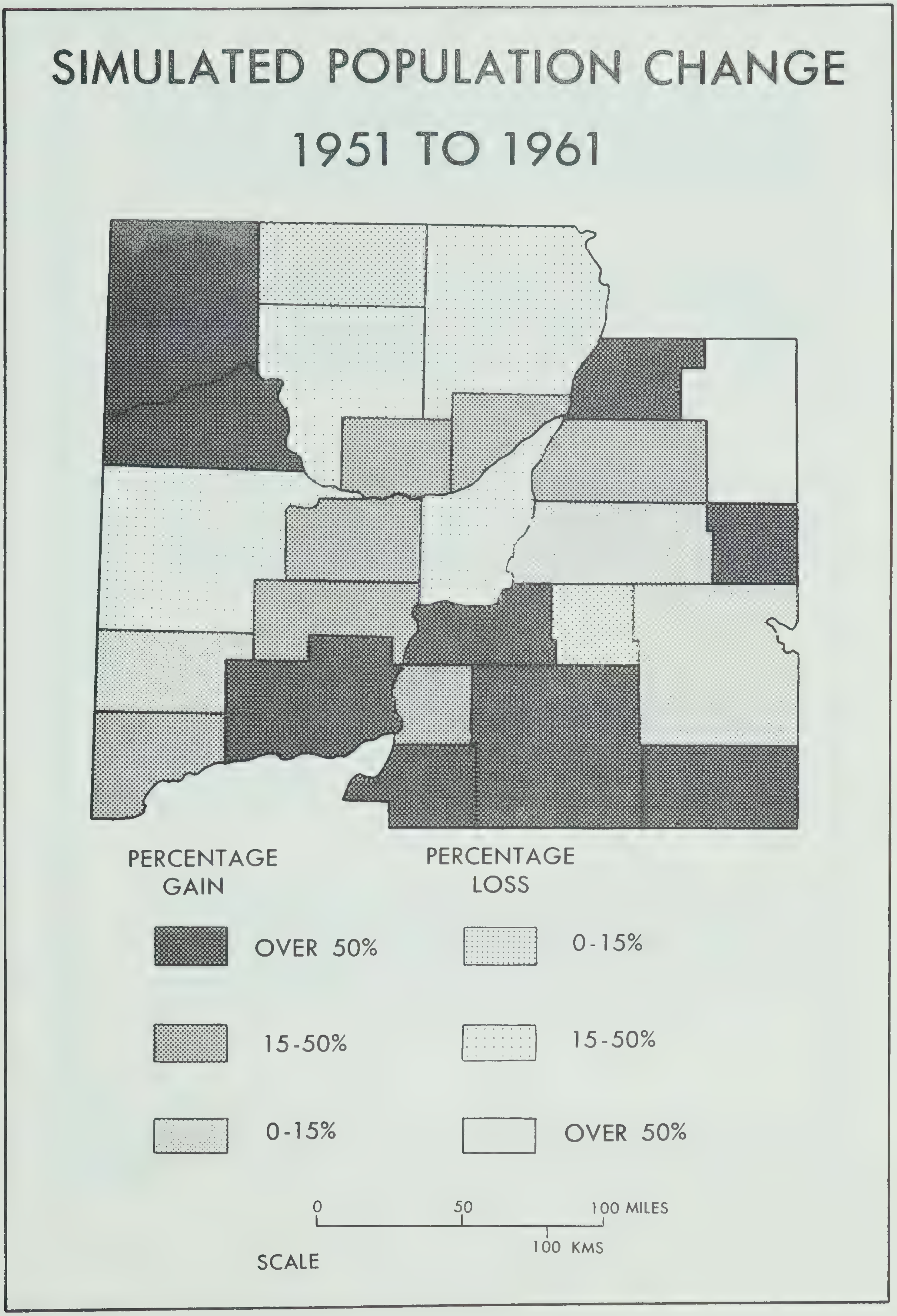

SIMULATED POPULATION CHANGE
1951 TO 1961
PERCENTAGE GAIN
OVER 50%
15-50%
0-15%
PERCENTAGE LOSS
0-15%
15-50%
OVER 50%
0
50
100 MILES
100 KMS
SCALE

FIGURE 21

TABLE 4.12 - TRANSITION MATRIX OF PROBABILITY OF MIGRATION BETWEEN REGIONS, 1921-1931

Region of Origin	Region of Destination				
	1	2	3	4	5
1	0.541	0.960×10^{-2}	0.932×10^{-4}	0.618×10^{-2}	0.243
2	0.166×10^{-4}	0.388	0.619×10^{-3}	0.504×10^{-1}	0.578×10^{-1}
3	0.159×10^{-5}	0.614×10^{-2}	0.542	0.761×10^{-1}	0.538×10^{-2}
4	0.101×10^{-5}	0.478×10^{-2}	0.729×10^{-3}	0.491	0.463×10^{-2}
5	0.552×10^{-4}	0.761×10^{-2}	0.715×10^{-4}	0.643×10^{-2}	0.453
6	0.329×10^{-5}	0.576×10^{-2}	0.460×10^{-4}	0.668×10^{-2}	0.468×10^{-1}
7	0.127×10^{-5}	0.452×10^{-2}	0.113×10^{-3}	0.450×10^{-1}	0.853×10^{-2}
8	0.872×10^{-5}	0.255×10^{-2}	0.416×10^{-4}	0.410×10^{-2}	0.208
9	0.467×10^{-6}	0.395×10^{-3}	0.617×10^{-5}	0.831×10^{-3}	0.757×10^{-2}
10	0.878×10^{-6}	0.368×10^{-3}	0.820×10^{-5}	0.839×10^{-3}	0.105×10^{-1}
11	0.476×10^{-6}	0.315×10^{-3}	0.681×10^{-5}	0.836×10^{-3}	0.596×10^{-2}
12	0.116×10^{-5}	0.685×10^{-3}	0.178×10^{-4}	0.201×10^{-2}	0.116×10^{-1}
13	0.256×10^{-6}	0.501×10^{-3}	0.113×10^{-3}	0.116×10^{-1}	0.111×10^{-2}
14	0.399×10^{-6}	0.460×10^{-3}	0.613×10^{-4}	0.361×10^{-2}	0.158×10^{-2}
15	0.690×10^{-7}	0.831×10^{-4}	0.975×10^{-5}	0.747×10^{-3}	0.300×10^{-3}
16	0.177×10^{-6}	0.311×10^{-3}	0.314×10^{-4}	0.574×10^{-2}	0.914×10^{-3}
17	0.255×10^{-6}	0.329×10^{-3}	0.102×10^{-4}	0.197×10^{-2}	0.182×10^{-2}
18	0.294×10^{-6}	0.435×10^{-3}	0.276×10^{-4}	0.487×10^{-2}	0.174×10^{-2}
19	0.329×10^{-6}	0.373×10^{-3}	0.308×10^{-4}	0.314×10^{-2}	0.163×10^{-2}
20	0.333×10^{-6}	0.326×10^{-3}	0.207×10^{-4}	0.218×10^{-2}	0.179×10^{-2}
21	0.307×10^{-6}	0.292×10^{-3}	0.951×10^{-5}	0.133×10^{-2}	0.254×10^{-2}
22	0.950×10^{-7}	0.110×10^{-3}	0.502×10^{-5}	0.747×10^{-3}	0.643×10^{-3}
23	0.400×10^{-7}	0.371×10^{-4}	0.178×10^{-5}	0.215×10^{-3}	0.251×10^{-3}
24	0.468×10^{-6}	0.340×10^{-3}	0.117×10^{-4}	0.135×10^{-2}	0.361×10^{-2}
25	0.396×10^{-6}	0.312×10^{-3}	0.141×10^{-4}	0.151×10^{-2}	0.248×10^{-2}
26	0.693×10^{-6}	0.578×10^{-3}	0.323×10^{-4}	0.317×10^{-2}	0.379×10^{-2}

TABLE 4.12 (Continued)

Region of Origin	Region of Destination 6	7	8	9	10
1	0.122×10^{-1}	0.536×10^{-5}	0.696×10^{-2}	0.379×10^{-1}	0.257×10^{-5}
2	0.369×10^{-1}	0.330×10^{-4}	0.346×10^{-2}	0.552×10^{-1}	0.186×10^{-5}
3	0.292×10^{-2}	0.815×10^{-5}	0.559×10^{-3}	0.856×10^{-2}	0.410×10^{-6}
4	0.406×10^{-2}	0.311×10^{-4}	0.528×10^{-3}	0.110×10^{-1}	0.402×10^{-6}
5	0.394×10^{-1}	0.820×10^{-5}	0.372×10^{-1}	0.140	0.696×10^{-5}
6	0.453	0.193×10^{-4}	0.399×10^{-2}	0.225	0.182×10^{-5}
7	0.169×10^{-1}	0.542	0.109×10^{-2}	0.396×10^{-1}	0.821×10^{-6}
8	0.189×10^{-1}	0.585×10^{-5}	0.100×10^{-3}	0.248	0.578×10^{-4}
9	0.103×10^{-1}	0.207×10^{-5}	0.240×10^{-2}	0.793	0.147×10^{-5}
10	0.230×10^{-2}	0.119×10^{-5}	0.155×10^{-1}	0.408×10^{-1}	0.542
11	0.310×10^{-2}	0.158×10^{-5}	0.385×10^{-2}	0.142	0.870×10^{-5}
12	0.440×10^{-2}	0.310×10^{-5}	0.743×10^{-2}	0.892×10^{-1}	0.465×10^{-4}
13	0.657×10^{-3}	0.253×10^{-5}	0.159×10^{-3}	0.270×10^{-2}	0.145×10^{-6}
14	0.665×10^{-3}	0.148×10^{-5}	0.250×10^{-3}	0.323×10^{-2}	0.248×10^{-6}
15	0.138×10^{-3}	0.339×10^{-6}	0.504×10^{-4}	0.709×10^{-3}	0.530×10^{-7}
16	0.630×10^{-3}	0.298×10^{-5}	0.150×10^{-3}	0.302×10^{-2}	0.152×10^{-6}
17	0.226×10^{-2}	0.699×10^{-5}	0.393×10^{-3}	0.197×10^{-1}	0.432×10^{-6}
18	0.127×10^{-2}	0.518×10^{-5}	0.334×10^{-3}	0.767×10^{-2}	0.380×10^{-6}
19	0.812×10^{-3}	0.201×10^{-5}	0.313×10^{-3}	0.481×10^{-2}	0.369×10^{-6}
20	0.853×10^{-3}	0.178×10^{-5}	0.396×10^{-3}	0.589×10^{-2}	0.528×10^{-6}
21	0.174×10^{-2}	0.245×10^{-5}	0.804×10^{-3}	0.247×10^{-1}	0.142×10^{-5}
22	0.439×10^{-3}	0.108×10^{-5}	0.154×10^{-3}	0.372×10^{-2}	0.215×10^{-6}
23	0.130×10^{-3}	0.243×10^{-6}	0.649×10^{-4}	0.112×10^{-2}	0.103×10^{-6}
24	0.161×10^{-2}	0.185×10^{-5}	0.136×10^{-2}	0.208×10^{-1}	0.349×10^{-5}
25	0.107×10^{-2}	0.162×10^{-5}	0.706×10^{-3}	0.971×10^{-2}	0.127×10^{-5}
26	0.159×10^{-2}	0.273×10^{-5}	0.908×10^{-3}	0.118×10^{-1}	0.134×10^{-5}

TABLE 4.12 (Continued)

Region of Origin	Region of Destination 11	12	13	14	15
1	0.474×10^{-2}	0.203×10^{-1}	0.332×10^{-1}	0.444×10^{-2}	0.653×10^{-2}
2	0.542×10^{-2}	0.207×10^{-1}	0.112	0.882×10^{-2}	0.135×10^{-1}
3	0.116×10^{-2}	0.543×10^{-2}	0.250	0.116×10^{-1}	0.157×10^{-1}
4	0.137×10^{-2}	0.578×10^{-2}	0.246	0.657×10^{-2}	0.116×10^{-1}
5	0.135×10^{-1}	0.462×10^{-1}	0.328×10^{-1}	0.400×10^{-2}	0.645×10^{-2}
6	0.832×10^{-2}	0.208×10^{-1}	0.229×10^{-1}	0.199×10^{-2}	0.352×10^{-2}
7	0.373×10^{-2}	0.128×10^{-1}	0.776×10^{-1}	0.388×10^{-2}	0.759×10^{-2}
8	0.488×10^{-1}	0.166	0.262×10^{-1}	0.354×10^{-2}	0.606×10^{-2}
9	0.174×10^{-1}	0.193×10^{-1}	0.431×10^{-2}	0.443×10^{-3}	0.826×10^{-3}
10	0.296×10^{-1}	0.278	0.643×10^{-2}	0.942×10^{-3}	0.170×10^{-2}
11	0.337	0.273	0.573×10^{-2}	0.693×10^{-3}	0.133×10^{-2}
12	0.156	0.171	0.166×10^{-1}	0.233×10^{-2}	0.453×10^{-2}
13	0.441×10^{-3}	0.225×10^{-2}	0.725	0.166×10^{-1}	0.358×10^{-1}
14	0.622×10^{-3}	0.367×10^{-2}	0.194	0.100×10^{-3}	0.627
15	0.141×10^{-3}	0.840×10^{-3}	0.491×10^{-1}	0.737×10^{-1}	0.755
16	0.528×10^{-3}	0.270×10^{-2}	0.263	0.591×10^{-2}	0.191×10^{-1}
17	0.304×10^{-2}	0.102×10^{-1}	0.122×10^{-1}	0.929×10^{-3}	0.206×10^{-2}
18	0.150×10^{-2}	0.788×10^{-2}	0.845×10^{-1}	0.524×10^{-2}	0.158×10^{-1}
19	0.106×10^{-2}	0.672×10^{-2}	0.117	0.297×10^{-1}	0.180
20	0.150×10^{-2}	0.107×10^{-1}	0.503×10^{-1}	0.116×10^{-1}	0.393×10^{-1}
21	0.108×10^{-1}	0.629×10^{-1}	0.118×10^{-1}	0.133×10^{-2}	0.296×10^{-2}
22	0.974×10^{-3}	0.580×10^{-2}	0.925×10^{-2}	0.947×10^{-3}	0.249×10^{-2}
23	0.348×10^{-3}	0.277×10^{-2}	0.321×10^{-2}	0.495×10^{-3}	0.131×10^{-2}
24	0.119×10^{-1}	0.201	0.141×10^{-1}	0.213×10^{-2}	0.457×10^{-2}
25	0.340×10^{-2}	0.337×10^{-1}	0.217×10^{-1}	0.415×10^{-2}	0.985×10^{-2}
26	0.335×10^{-2}	0.273×10^{-1}	0.602×10^{-1}	0.160×10^{-1}	0.422×10^{-1}

TABLE 4.12 (Continued)

Region of Origin	Region of Destination 16	17	18	19	20
1	0.438×10^{-2}	0.771×10^{-2}	0.143×10^{-1}	0.543×10^{-3}	0.578×10^{-6}
2	0.133×10^{-1}	0.195×10^{-1}	0.366×10^{-1}	0.106×10^{-2}	0.977×10^{-6}
3	0.133×10^{-1}	0.599×10^{-2}	0.230×10^{-1}	0.869×10^{-3}	0.615×10^{-6}
4	0.232×10^{-1}	0.111×10^{-1}	0.389×10^{-1}	0.850×10^{-3}	0.621×10^{-6}
5	0.514×10^{-2}	0.142×10^{-1}	0.192×10^{-1}	0.614×10^{-3}	0.709×10^{-6}
6	0.420×10^{-2}	0.209×10^{-1}	0.167×10^{-1}	0.361×10^{-3}	0.399×10^{-6}
7	0.176×10^{-1}	0.568×10^{-1}	0.597×10^{-1}	0.786×10^{-3}	0.732×10^{-6}
8	0.474×10^{-2}	0.172×10^{-1}	0.207×10^{-1}	0.659×10^{-3}	0.871×10^{-6}
9	0.924×10^{-3}	0.833×10^{-2}	0.460×10^{-2}	0.980×10^{-4}	0.126×10^{-6}
10	0.128×10^{-2}	0.507×10^{-2}	0.632×10^{-2}	0.208×10^{-3}	0.313×10^{-6}
11	0.131×10^{-2}	0.105×10^{-1}	0.735×10^{-2}	0.175×10^{-3}	0.261×10^{-6}
12	0.380×10^{-2}	0.201×10^{-1}	0.219×10^{-1}	0.633×10^{-3}	0.106×10^{-5}
13	0.501×10^{-1}	0.323×10^{-2}	0.318×10^{-1}	0.150×10^{-2}	0.675×10^{-6}
14	0.131×10^{-1}	0.287×10^{-2}	0.229×10^{-1}	0.441×10^{-2}	0.181×10^{-5}
15	0.499×10^{-2}	0.748×10^{-3}	0.817×10^{-2}	0.315×10^{-2}	0.723×10^{-6}
16	0.340	0.603×10^{-2}	0.207	0.240×10^{-2}	0.112×10^{-5}
17	0.434×10^{-2}	0.604	0.372×10^{-1}	0.308×10^{-3}	0.388×10^{-6}
18	0.105	0.263×10^{-1}	0.226	0.379×10^{-2}	0.330×10^{-5}
19	0.359×10^{-1}	0.641×10^{-2}	0.112	0.302	0.313×10^{-4}
20	0.159×10^{-1}	0.768×10^{-2}	0.926×10^{-1}	0.298×10^{-1}	0.542
21	0.370×10^{-2}	0.486×10^{-1}	0.325×10^{-1}	0.517×10^{-3}	0.924×10^{-6}
22	0.434×10^{-2}	0.143×10^{-1}	0.841×10^{-1}	0.610×10^{-3}	0.116×10^{-5}
23	0.112×10^{-2}	0.177×10^{-2}	0.109×10^{-1}	0.389×10^{-3}	0.166×10^{-5}
24	0.373×10^{-2}	0.149×10^{-1}	0.260×10^{-1}	0.800×10^{-3}	0.176×10^{-5}
25	0.597×10^{-2}	0.970×10^{-2}	0.411×10^{-1}	0.231×10^{-2}	0.919×10^{-5}
26	0.154×10^{-1}	0.121×10^{-1}	0.839×10^{-1}	0.130×10^{-1}	0.109×10^{-3}

TABLE 4.12 (Continued)

Region of Origin	Region of Destination 21	22	23	24	25	26
1	0.257×10^{-1}	0.196×10^{-1}	0.152×10^{-3}	0.101×10^{-2}	0.548×10^{-6}	0.393×10^{-6}
2	0.422×10^{-1}	0.389×10^{-1}	0.246×10^{-3}	0.127×10^{-2}	0.745×10^{-6}	0.565×10^{-6}
3	0.136×10^{-1}	0.176×10^{-1}	0.117×10^{-3}	0.432×10^{-3}	0.333×10^{-6}	0.313×10^{-6}
4	0.182×10^{-1}	0.251×10^{-1}	0.135×10^{-3}	0.478×10^{-3}	0.343×10^{-6}	0.294×10^{-6}
5	0.484×10^{-1}	0.300×10^{-1}	0.219×10^{-3}	0.177×10^{-2}	0.781×10^{-6}	0.488×10^{-6}
6	0.393×10^{-1}	0.243×10^{-1}	0.134×10^{-3}	0.941×10^{-3}	0.400×10^{-6}	0.242×10^{-6}
7	0.486×10^{-1}	0.525×10^{-1}	0.220×10^{-3}	0.947×10^{-3}	0.530×10^{-6}	0.366×10^{-6}
8	0.857×10^{-1}	0.403×10^{-1}	0.317×10^{-3}	0.374×10^{-2}	0.124×10^{-5}	0.655×10^{-6}
9	0.255×10^{-1}	0.941×10^{-2}	0.531×10^{-4}	0.553×10^{-3}	0.166×10^{-6}	0.830×10^{-7}
10	0.405×10^{-1}	0.151×10^{-1}	0.135×10^{-3}	0.258×10^{-2}	0.600×10^{-6}	0.259×10^{-6}
11	0.904×10^{-1}	0.201×10^{-1}	0.134×10^{-3}	0.258×10^{-2}	0.472×10^{-6}	0.190×10^{-6}
12	0.301	0.680×10^{-1}	0.607×10^{-3}	0.248×10^{-1}	0.267×10^{-5}	0.884×10^{-6}
13	0.762×10^{-2}	0.147×10^{-1}	0.950×10^{-4}	0.235×10^{-3}	0.232×10^{-6}	0.263×10^{-6}
14	0.100×10^{-1}	0.175×10^{-1}	0.171×10^{-3}	0.415×10^{-3}	0.517×10^{-6}	0.814×10^{-6}
15	0.263×10^{-2}	0.541×10^{-2}	0.531×10^{-4}	0.105×10^{-3}	0.144×10^{-6}	0.253×10^{-6}
16	0.125×10^{-1}	0.360×10^{-1}	0.174×10^{-3}	0.327×10^{-3}	0.334×10^{-6}	0.353×10^{-6}
17	0.118	0.857×10^{-1}	0.198×10^{-3}	0.940×10^{-3}	0.391×10^{-6}	0.199×10^{-6}
18	0.560×10^{-1}	0.355	0.861×10^{-3}	0.115×10^{-2}	0.117×10^{-5}	0.977×10^{-6}
19	0.262×10^{-1}	0.758×10^{-1}	0.902×10^{-3}	0.105×10^{-2}	0.194×10^{-5}	0.447×10^{-5}
20	0.446×10^{-1}	0.137	0.367×10^{-2}	0.219×10^{-2}	0.733×10^{-5}	0.356×10^{-4}
21	0.453	0.233	0.917×10^{-3}	0.122×10^{-1}	0.200×10^{-5}	0.597×10^{-6}
22	0.952×10^{-1}	0.680	0.141×10^{-2}	0.130×10^{-2}	0.109×10^{-5}	0.482×10^{-6}
23	0.200×10^{-1}	0.756×10^{-1}	0.785	0.104×10^{-2}	0.443×10^{-5}	0.100×10^{-5}
24	0.472	0.124	0.185×10^{-2}	0.100×10^{-3}	0.104×10^{-4}	0.165×10^{-5}
25	0.121	0.161	0.123×10^{-1}	0.162×10^{-1}	0.542	0.163×10^{-4}
26	0.882×10^{-1}	0.175	0.680×10^{-2}	0.632×10^{-2}	0.398×10^{-4}	0.428

TABLE 4.13 - TRANSITION MATRIX OF PROBABILITY OF MIGRATION BETWEEN REGIONS, 1931-1941

Region of Origin	Region of Destination				
	1	2	3	4	5
1	0.516	0.133×10^{-1}	0.580×10^{-3}	0.128×10^{-1}	0.197
2	0.861×10^{-3}	0.404	0.256×10^{-2}	0.676×10^{-1}	0.466×10^{-1}
3	0.121×10^{-3}	0.828×10^{-2}	0.495	0.107	0.638×10^{-2}
4	0.638×10^{-4}	0.521×10^{-2}	0.255×10^{-2}	0.622	0.438×10^{-2}
5	0.493×10^{-2}	0.180×10^{-1}	0.762×10^{-3}	0.220×10^{-1}	0.100×10^{-3}
6	0.232×10^{-3}	0.792×10^{-2}	0.290×10^{-3}	0.127×10^{-1}	0.431×10^{-1}
7	0.768×10^{-4}	0.489×10^{-2}	0.488×10^{-3}	0.521×10^{-1}	0.739×10^{-2}
8	0.599×10^{-3}	0.425×10^{-2}	0.292×10^{-3}	0.912×10^{-2}	0.176
9	0.120×10^{-3}	0.216×10^{-2}	0.143×10^{-3}	0.588×10^{-2}	0.250×10^{-1}
10	0.355×10^{-4}	0.345×10^{-3}	0.312×10^{-4}	0.101×10^{-2}	0.565×10^{-2}
11	0.584×10^{-4}	0.849×10^{-3}	0.747×10^{-4}	0.283×10^{-2}	0.971×10^{-2}
12	0.898×10^{-4}	0.118×10^{-2}	0.122×10^{-3}	0.431×10^{-2}	0.123×10^{-1}
13	0.266×10^{-4}	0.100×10^{-2}	0.690×10^{-3}	0.224×10^{-1}	0.175×10^{-2}
14	0.372×10^{-4}	0.880×10^{-3}	0.383×10^{-3}	0.760×10^{-2}	0.225×10^{-2}
15	0.214×10^{-4}	0.527×10^{-3}	0.205×10^{-3}	0.513×10^{-2}	0.141×10^{-2}
16	0.221×10^{-4}	0.758×10^{-3}	0.258×10^{-3}	0.139×10^{-1}	0.169×10^{-2}
17	0.242×10^{-4}	0.706×10^{-3}	0.853×10^{-4}	0.482×10^{-2}	0.274×10^{-2}
18	0.202×10^{-4}	0.594×10^{-3}	0.134×10^{-3}	0.702×10^{-2}	0.173×10^{-2}
19	0.235×10^{-4}	0.549×10^{-3}	0.156×10^{-3}	0.504×10^{-2}	0.174×10^{-2}
20	0.285×10^{-4}	0.585×10^{-3}	0.133×10^{-3}	0.440×10^{-2}	0.226×10^{-2}
21	0.331×10^{-4}	0.660×10^{-3}	0.832×10^{-4}	0.353×10^{-2}	0.382×10^{-2}
22	0.214×10^{-4}	0.504×10^{-3}	0.856×10^{-4}	0.385×10^{-2}	0.206×10^{-2}
23	0.514×10^{-5}	0.101×10^{-3}	0.179×10^{-4}	0.669×10^{-3}	0.469×10^{-3}
24	0.621×10^{-4}	0.977×10^{-3}	0.129×10^{-3}	0.464×10^{-2}	0.674×10^{-2}
25	0.308×10^{-4}	0.521×10^{-3}	0.875×10^{-4}	0.295×10^{-2}	0.279×10^{-2}
26	0.523×10^{-4}	0.931×10^{-3}	0.189×10^{-3}	0.587×10^{-2}	0.420×10^{-2}

TABLE 4.13 (Continued)

Region of Origin	Region of Destination 6	7	8	9	10
1	0.366×10^{-1}	0.576×10^{-3}	0.653×10^{-2}	0.541×10^{-1}	0.716×10^{-4}
2	0.806×10^{-1}	0.236×10^{-2}	0.300×10^{-2}	0.630×10^{-1}	0.450×10^{-4}
3	0.957×10^{-2}	0.764×10^{-3}	0.666×10^{-3}	0.135×10^{-1}	0.132×10^{-4}
4	0.100×10^{-1}	0.194×10^{-2}	0.496×10^{-3}	0.132×10^{-1}	0.101×10^{-4}
5	0.170	0.138×10^{-2}	0.480×10^{-1}	0.282	0.285×10^{-3}
6	0.477	0.165×10^{-2}	0.379×10^{-2}	0.241	0.492×10^{-4}
7	0.347×10^{-1}	0.408	0.923×10^{-3}	0.401×10^{-1}	0.187×10^{-4}
8	0.547×10^{-1}	0.633×10^{-3}	0.100×10^{-3}	0.288	0.113×10^{-2}
9	0.844×10^{-1}	0.665×10^{-3}	0.698×10^{-2}	0.505	0.118×10^{-3}
10	0.384×10^{-2}	0.694×10^{-4}	0.610×10^{-2}	0.263×10^{-1}	0.180
11	0.140×10^{-1}	0.252×10^{-3}	0.505×10^{-2}	0.221	0.268×10^{-3}
12	0.134×10^{-1}	0.320×10^{-3}	0.633×10^{-2}	0.103	0.821×10^{-3}
13	0.282×10^{-2}	0.299×10^{-3}	0.241×10^{-3}	0.535×10^{-2}	0.577×10^{-5}
14	0.270×10^{-2}	0.176×10^{-3}	0.340×10^{-3}	0.593×10^{-2}	0.875×10^{-5}
15	0.183×10^{-2}	0.130×10^{-3}	0.224×10^{-3}	0.421×10^{-2}	0.604×10^{-5}
16	0.313×10^{-2}	0.397×10^{-3}	0.265×10^{-3}	0.682×10^{-2}	0.692×10^{-5}
17	0.853×10^{-2}	0.745×10^{-3}	0.546×10^{-3}	0.313×10^{-1}	0.154×10^{-4}
18	0.339×10^{-2}	0.376×10^{-3}	0.311×10^{-3}	0.898×10^{-2}	0.903×10^{-5}
19	0.241×10^{-2}	0.173×10^{-3}	0.311×10^{-3}	0.631×10^{-2}	0.929×10^{-5}
20	0.302×10^{-2}	0.187×10^{-3}	0.456×10^{-3}	0.904×10^{-2}	0.153×10^{-4}
21	0.704×10^{-2}	0.308×10^{-3}	0.106×10^{-2}	0.396×10^{-1}	0.452×10^{-4}
22	0.378×10^{-2}	0.270×10^{-3}	0.448×10^{-3}	0.135×10^{-1}	0.155×10^{-4}
23	0.676×10^{-3}	0.379×10^{-4}	0.109×10^{-3}	0.246×10^{-2}	0.424×10^{-5}
24	0.854×10^{-2}	0.312×10^{-3}	0.218×10^{-2}	0.441×10^{-1}	0.129×10^{-3}
25	0.342×10^{-2}	0.159×10^{-3}	0.705×10^{-3}	0.130×10^{-1}	0.305×10^{-4}
26	0.502×10^{-2}	0.262×10^{-3}	0.914×10^{-3}	0.161×10^{-1}	0.333×10^{-4}

TABLE 4.13 (Continued)

Region of Origin	Region of Destination				
	11	12	13	14	15
1	0.113×10^{-1}	0.128×10^{-1}	0.341×10^{-1}	0.684×10^{-2}	0.118×10^{-1}
2	0.106×10^{-1}	0.109×10^{-1}	0.830×10^{-1}	0.104×10^{-1}	0.188×10^{-1}
3	0.303×10^{-2}	0.366×10^{-2}	0.185	0.147×10^{-1}	0.237×10^{-1}
4	0.274×10^{-2}	0.307×10^{-2}	0.143	0.696×10^{-2}	0.141×10^{-1}
5	0.471×10^{-1}	0.439×10^{-1}	0.559×10^{-1}	0.104×10^{-1}	0.195×10^{-1}
6	0.173×10^{-1}	0.122×10^{-1}	0.229×10^{-1}	0.315×10^{-2}	0.642×10^{-2}
7	0.654×10^{-2}	0.612×10^{-2}	0.512×10^{-1}	0.433×10^{-2}	0.963×10^{-2}
8	0.898×10^{-1}	0.829×10^{-1}	0.283×10^{-1}	0.572×10^{-2}	0.113×10^{-1}
9	0.952×10^{-1}	0.328×10^{-1}	0.152×10^{-1}	0.242×10^{-2}	0.516×10^{-2}
10	0.258×10^{-1}	0.582×10^{-1}	0.366×10^{-2}	0.797×10^{-3}	0.165×10^{-2}
11	0.349	0.162	0.935×10^{-2}	0.172×10^{-2}	0.376×10^{-2}
12	0.219	0.170	0.167×10^{-1}	0.350×10^{-2}	0.775×10^{-2}
13	0.142×10^{-2}	0.187×10^{-2}	0.669	0.220×10^{-1}	0.533×10^{-1}
14	0.181×10^{-2}	0.273×10^{-2}	0.153	0.100×10^{-3}	0.623
15	0.132×10^{-2}	0.201×10^{-2}	0.123	0.207	0.459
16	0.191×10^{-2}	0.252×10^{-2}	0.243	0.102×10^{-1}	0.352×10^{-1}
17	0.789×10^{-2}	0.724×10^{-2}	0.145×10^{-1}	0.178×10^{-2}	0.442×10^{-2}
18	0.279×10^{-2}	0.378×10^{-2}	0.524×10^{-1}	0.535×10^{-2}	0.175×10^{-1}
19	0.217×10^{-2}	0.347×10^{-2}	0.740×10^{-1}	0.260×10^{-1}	0.156
20	0.353×10^{-2}	0.628×10^{-2}	0.422×10^{-1}	0.136×10^{-1}	0.492×10^{-1}
21	0.248×10^{-1}	0.369×10^{-1}	0.146×10^{-1}	0.254×10^{-2}	0.632×10^{-2}
22	0.543×10^{-2}	0.821×10^{-1}	0.213×10^{-1}	0.339×10^{-2}	0.977×10^{-2}
23	0.114×10^{-2}	0.224×10^{-1}	0.438×10^{-2}	0.998×10^{-3}	0.289×10^{-2}
24	0.350×10^{-1}	0.133×	0.222×10^{-1}	0.497×10^{-2}	0.120×10^{-1}
25	0.672×10^{-2}	0.159×10^{-1}	0.186×10^{-1}	0.512×10^{-2}	0.135×10^{-1}
26	0.689×10^{-2}	0.137×10^{-1}	0.474×10^{-1}	0.174×10^{-1}	0.503×10^{-1}

TABLE 4.13 (Continued)

Region of Origin	Region of Destination 16	17	18	19	20
1	0.100×10^{-1}	0.700×10^{-2}	0.162×10^{-1}	0.204×10^{-2}	0.733×10^{-4}
2	0.222×10^{-1}	0.132×10^{-1}	0.309×10^{-1}	0.309×10^{-2}	0.972×10^{-4}
3	0.245×10^{-1}	0.516×10^{-2}	0.227×10^{-1}	0.285×10^{-2}	0.713×10^{-4}
4	0.314×10^{-1}	0.694×10^{-2}	0.282×10^{-1}	0.219×10^{-2}	0.563×10^{-4}
5	0.192×10^{-1}	0.198×10^{-1}	0.350×10^{-1}	0.378×10^{-2}	0.146×10^{-3}
6	0.901×10^{-2}	0.157×10^{-1}	0.173×10^{-1}	0.133×10^{-2}	0.494×10^{-4}
7	0.240×10^{-1}	0.288×10^{-1}	0.405×10^{-1}	0.202×10^{-2}	0.643×10^{-4}
8	0.110×10^{-1}	0.145×10^{-1}	0.229×10^{-1}	0.248×10^{-2}	0.107×10^{-3}
9	0.686×10^{-2}	0.201×10^{-1}	0.161×10^{-1}	0.122×10^{-2}	0.515×10^{-4}
10	0.156×10^{-2}	0.220×10^{-2}	0.361×10^{-2}	0.401×10^{-3}	0.195×10^{-4}
11	0.446×10^{-2}	0.118×10^{-1}	0.116×10^{-1}	0.973×10^{-3}	0.468×10^{-4}
12	0.800×10^{-2}	0.146×10^{-1}	0.213×10^{-1}	0.211×10^{-2}	0.113×10^{-3}
13	0.861×10^{-1}	0.328×10^{-2}	0.330×10^{-1}	0.502×10^{-2}	0.846×10^{-4}
14	0.252×10^{-1}	0.280×10^{-2}	0.234×10^{-1}	0.123×10^{-1}	0.191×10^{-3}
15	0.289×10^{-1}	0.231×10^{-2}	0.255×10^{-1}	0.246×10^{-1}	0.229×10^{-3}
16	0.356	0.653×10^{-2}	0.197	0.876×10^{-2}	0.152×10^{-3}
17	0.102×10^{-1}	0.627	0.387×10^{-1}	0.128×10^{-2}	0.532×10^{-4}
18	0.111	0.139×10^{-1}	0.457	0.763×10^{-2}	0.229×10^{-3}
19	0.456×10^{-1}	0.425×10^{-2}	0.707×10^{-1}	0.480	0.175×10^{-2}
20	0.268×10^{-1}	0.598×10^{-2}	0.718×10^{-1}	0.592×10^{-1}	0.516
21	0.923×10^{-2}	0.376×10^{-1}	0.356×10^{-1}	0.209×10^{-2}	0.118×10^{-3}
22	0.192×10^{-1}	0.232×10^{-1}	0.148	0.436×10^{-2}	0.260×10^{-3}
23	0.304×10^{-2}	0.193×10^{-2}	0.128×10^{-1}	0.153×10^{-2}	0.186×10^{-3}
24	0.121×10^{-1}	0.173×10^{-1}	0.379×10^{-1}	0.397×10^{-2}	0.270×10^{-3}
25	0.105×10^{-1}	0.680×10^{-2}	0.326×10^{-1}	0.580×10^{-2}	0.662×10^{-3}
26	0.249×10^{-1}	0.855×10^{-2}	0.634×10^{-1}	0.275×10^{-1}	0.605×10^{-2}

TABLE 4.13 (continued)

Region of Origin	Region of Destination					
	21	22	23	24	25	26
1	0.240×10^{-1}	0.205×10^{-1}	0.316×10^{-2}	0.275×10^{-2}	0.516×10^{-4}	0.129×10^{-3}
2	0.309×10^{-1}	0.312×10^{-1}	0.403×10^{-2}	0.280×10^{-2}	0.565×10^{-4}	0.149×10^{-3}
3	0.126×10^{-1}	0.172×10^{-1}	0.231×10^{-2}	0.120×10^{-2}	0.307×10^{-4}	0.976×10^{-4}
4	0.128×10^{-1}	0.184×10^{-1}	0.205×10^{-2}	0.102×10^{-2}	0.246×10^{-4}	0.723×10^{-4}
5	0.693×10^{-1}	0.494×10^{-1}	0.722×10^{-2}	0.747×10^{-2}	0.117×10^{-3}	0.259×10^{-3}
6	0.324×10^{-1}	0.230×10^{-1}	0.265×10^{-2}	0.240×10^{-2}	0.365×10^{-4}	0.788×10^{-4}
7	0.298×10^{-1}	0.346×10^{-1}	0.312×10^{-2}	0.185×10^{-2}	0.357×10^{-4}	0.865×10^{-4}
8	0.704×10^{-1}	0.393×10^{-1}	0.617×10^{-2}	0.885×10^{-2}	0.108×10^{-3}	0.207×10^{-3}
9	0.637×10^{-1}	0.288×10^{-1}	0.336×10^{-2}	0.433×10^{-2}	0.483×10^{-4}	0.881×10^{-4}
10	0.162×10^{-1}	0.739×10^{-2}	0.130×10^{-2}	0.285×10^{-2}	0.254×10^{-4}	0.408×10^{-4}
11	0.926×10^{-1}	0.268×10^{-1}	0.363×10^{-2}	0.801×10^{-2}	0.581×10^{-4}	0.877×10^{-4}
12	0.187	0.550×10^{-1}	0.964×10^{-2}	0.413×10^{-1}	0.187×10^{-3}	0.238×10^{-3}
13	0.828×10^{-2}	0.160×10^{-1}	0.211×10^{-2}	0.769×10^{-3}	0.244×10^{-4}	0.915×10^{-4}
14	0.100×10^{-1}	0.177×10^{-1}	0.334×10^{-2}	0.120×10^{-2}	0.467×10^{-4}	0.234×10^{-3}
15	0.828×10^{-2}	0.169×10^{-1}	0.322×10^{-2}	0.960×10^{-3}	0.410×10^{-4}	0.225×10^{-3}
16	0.147×10^{-1}	0.405×10^{-1}	0.413×10^{-2}	0.118×10^{-2}	0.387×10^{-4}	0.136×10^{-3}
17	0.942×10^{-1}	0.769×10^{-1}	0.410×10^{-2}	0.265×10^{-2}	0.394×10^{-4}	0.731×10^{-4}
18	0.320×10^{-1}	0.176	0.981×10^{-2}	0.208×10^{-2}	0.678×10^{-4}	0.194×10^{-3}
19	0.174×10^{-1}	0.480×10^{-1}	0.108×10^{-1}	0.202×10^{-2}	0.112×10^{-3}	0.782×10^{-3}
20	0.333×10^{-1}	0.968×10^{-1}	0.445×10^{-1}	0.465×10^{-2}	0.432×10^{-3}	0.581×10^{-2}
21	0.461	0.192	0.163×10^{-1}	0.261×10^{-1}	0.171×10^{-3}	0.199×10^{-3}
22	0.145	0.463	0.431×10^{-1}	0.660×10^{-2}	0.181×10^{-3}	0.297×10^{-3}
23	0.192×10^{-1}	0.671×10^{-1}	0.815	0.282×10^{-2}	0.324×10^{-3}	0.295×10^{-3}
24	0.426	0.143	0.392×10^{-1}	0.100×10^{-3}	0.946×10^{-3}	0.632×10^{-3}
25	0.739×10^{-1}	0.104	0.119	0.250×10^{-1}	0.494	0.271×10^{-2}
26	0.583×10^{-1}	0.115	0.736×10^{-1}	0.113×10^{-1}	0.184×10^{-2}	0.420

TABLE 4.14 - TRANSITION MATRIX OF PROBABILITY OF MIGRATION BETWEEN REGIONS, 1941-1951

Region of Origin	Region of Destination				
	1	2	3	4	5
1	0.415	0.196×10^{-1}	0.150×10^{-2}	0.157×10^{-1}	0.204
2	0.714×10^{-3}	0.481	0.694×10^{-2}	0.889×10^{-1}	0.350×10^{-1}
3	0.344×10^{-4}	0.438×10^{-2}	0.161	0.654×10^{-1}	0.164×10^{-2}
4	0.843×10^{-4}	0.131×10^{-1}	0.153×10^{-1}	0.451	0.541×10^{-2}
5	0.496×10^{-2}	0.234×10^{-1}	0.173×10^{-2}	0.245×10^{-1}	0.131
6	0.389×10^{-3}	0.227×10^{-1}	0.143×10^{-2}	0.324×10^{-1}	0.766×10^{-1}
7	0.470×10^{-4}	0.553×10^{-2}	0.107×10^{-2}	0.664×10^{-1}	0.441×10^{-2}
8	0.533×10^{-3}	0.531×10^{-2}	0.677×10^{-3}	0.105×10^{-1}	0.176
9	0.156×10^{-3}	0.442×10^{-2}	0.541×10^{-3}	0.114×10^{-1}	0.350×10^{-1}
10	0.895×10^{-4}	0.127×10^{-2}	0.221×10^{-3}	0.355×10^{-2}	0.148×10^{-1}
11	0.387×10^{-4}	0.863×10^{-3}	0.146×10^{-3}	0.280×10^{-2}	0.673×10^{-2}
12	0.636×10^{-4}	0.127×10^{-2}	0.257×10^{-3}	0.455×10^{-2}	0.887×10^{-2}
13	0.686×10^{-5}	0.447×10^{-3}	0.765×10^{-3}	0.124×10^{-1}	0.418×10^{-3}
14	0.151×10^{-4}	0.584×10^{-2}	0.595×10^{-3}	0.554×10^{-1}	0.843×10^{-3}
15	0.939×10^{-5}	0.378×10^{-3}	0.339×10^{-3}	0.410×10^{-2}	0.573×10^{-3}
16	0.151×10^{-4}	0.883×10^{-3}	0.683×10^{-3}	0.195×10^{-1}	0.109×10^{-2}
17	0.359×10^{-4}	0.175×10^{-2}	0.423×10^{-3}	0.128×10^{-1}	0.405×10^{-2}
18	0.907×10^{-5}	0.448×10^{-3}	0.219×10^{-3}	0.604×10^{-2}	0.750×10^{-3}
19	0.988×10^{-5}	0.376×10^{-3}	0.238×10^{-3}	0.382×10^{-2}	0.689×10^{-3}
20	0.204×10^{-4}	0.670×10^{-3}	0.328×10^{-3}	0.543×10^{-2}	0.154×10^{-2}
21	0.240×10^{-4}	0.764×10^{-3}	0.193×10^{-3}	0.423×10^{-3}	0.277×10^{-2}
22	0.287×10^{-4}	0.110×10^{-2}	0.391×10^{-3}	0.911×10^{-2}	0.270×10^{-2}
23	0.178×10^{-4}	0.558×10^{-3}	0.207×10^{-3}	0.392×10^{-2}	0.157×10^{-2}
24	0.443×10^{-4}	0.108×10^{-2}	0.289×10^{-3}	0.523×10^{-2}	0.476×10^{-2}
25	0.226×10^{-4}	0.597×10^{-3}	0.209×10^{-3}	0.352×10^{-2}	0.198×10^{-2}
26	0.171×10^{-4}	0.479×10^{-3}	0.207×10^{-3}	0.319×10^{-2}	0.131×10^{-2}

TABLE 4.14 (Continued)

Region of Origin	Region of Destination 6	7	8	9	10
1	0.342×10^{-1}	0.191×10^{-2}	0.697×10^{-2}	0.457×10^{-1}	0.934×10^{-6}
2	0.724×10^{-1}	0.821×10^{-2}	0.253×10^{-2}	0.472×10^{-1}	0.482×10^{-6}
3	0.288×10^{-2}	0.101×10^{-2}	0.204×10^{-3}	0.365×10^{-2}	0.530×10^{-7}
4	0.153×10^{-1}	0.145×10^{-1}	0.737×10^{-3}	0.180×10^{-1}	0.199×10^{-6}
5	0.164	0.438×10^{-2}	0.559×10^{-1}	0.250	0.376×10^{-5}
6	0.100×10^{-3}	0.130×10^{-1}	0.784×10^{-2}	0.511	0.127×10^{-5}
7	0.281×10^{-1}	0.280	0.671×10^{-3}	0.284×10^{-1}	0.180×10^{-6}
8	0.527×10^{-1}	0.209×10^{-2}	0.100×10^{-3}	0.295	0.204×10^{-4}
9	0.153	0.395×10^{-2}	0.131×10^{-1}	0.262	0.286×10^{-5}
10	0.107×10^{-1}	0.702×10^{-3}	0.255×10^{-1}	0.803×10^{-1}	0.484
11	0.114×10^{-1}	0.740×10^{-3}	0.510×10^{-2}	0.218	0.404×10^{-5}
12	0.109×10^{-1}	0.981×10^{-3}	0.665×10^{-2}	0.939×10^{-1}	0.144×10^{-4}
13	0.801×10^{-3}	0.386×10^{-3}	0.713×10^{-4}	0.142×10^{-2}	0.230×10^{-7}
14	0.115×10^{-2}	0.321×10^{-3}	0.159×10^{-3}	0.241×10^{-2}	0.560×10^{-7}
15	0.858×10^{-3}	0.264×10^{-3}	0.114×10^{-3}	0.189×10^{-2}	0.420×10^{-7}
16	0.243×10^{-2}	0.143×10^{-2}	0.214×10^{-3}	0.503×10^{-2}	0.760×10^{-7}
17	0.162×10^{-1}	0.627×10^{-2}	0.104×10^{-2}	0.603×10^{-1}	0.404×10^{-6}
18	0.178×10^{-2}	0.902×10^{-3}	0.171×10^{-3}	0.459×10^{-2}	0.690×10^{-7}
19	0.111×10^{-2}	0.346×10^{-3}	0.157×10^{-3}	0.283×10^{-2}	0.650×10^{-7}
20	0.238×10^{-2}	0.625×10^{-3}	0.401×10^{-3}	0.703×10^{-2}	0.189×10^{-6}
21	0.613×10^{-2}	0.109×10^{-2}	0.103×10^{-2}	0.369×10^{-1}	0.639×10^{-6}
22	0.597×10^{-2}	0.184×10^{-2}	0.767×10^{-3}	0.216×10^{-1}	0.376×10^{-6}
23	0.265×10^{-2}	0.621×10^{-3}	0.483×10^{-3}	0.977×10^{-2}	0.269×10^{-6}
24	0.693×10^{-2}	0.100×10^{-2}	0.211×10^{-2}	0.379×10^{-1}	0.190×10^{-5}
25	0.278×10^{-2}	0.529×10^{-3}	0.666×10^{-3}	0.107×10^{-1}	0.419×10^{-6}
26	0.179×10^{-2}	0.387×10^{-3}	0.372×10^{-3}	0.569×10^{-2}	0.193×10^{-6}

TABLE 4.14 (Continued)

Region of Origin	Region of Destination 11	12	13	14	15
1	0.856×10^{-2}	0.873×10^{-2}	0.218×10^{-1}	0.405×10^{-2}	0.708×10^{-2}
2	0.695×10^{-2}	0.633×10^{-2}	0.518×10^{-1}	0.569×10^{-2}	0.104×10^{-1}
3	0.741×10^{-3}	0.811×10^{-3}	0.560×10^{-1}	0.367×10^{-2}	0.589×10^{-2}
4	0.333×10^{-2}	0.336×10^{-2}	0.212	0.797×10^{-2}	0.166×10^{-1}
5	0.362×10^{-1}	0.296×10^{-1}	0.324×10^{-1}	0.550×10^{-2}	0.105×10^{-1}
6	0.286×10^{-1}	0.171×10^{-1}	0.290×10^{-1}	0.352×10^{-2}	0.737×10^{-2}
7	0.402×10^{-2}	0.330×10^{-2}	0.301×10^{-1}	0.211×10^{-2}	0.488×10^{-2}
8	0.863×10^{-1}	0.699×10^{-1}	0.173×10^{-1}	0.325×10^{-2}	0.660×10^{-2}
9	0.164	0.440×10^{-1}	0.154×10^{-1}	0.220×10^{-2}	0.486×10^{-2}
10	0.856×10^{-1}	0.190	0.702×10^{-2}	0.143×10^{-2}	0.305×10^{-2}
11	0.395	0.148	0.497×10^{-2}	0.836×10^{-3}	0.190×10^{-2}
12	0.238	0.226	0.969×10^{-2}	0.189×10^{-2}	0.434×10^{-2}
13	0.346×10^{-3}	0.418×10^{-3}	0.820	0.633×10^{-2}	0.161×10^{-1}
14	0.690×10^{-3}	0.968×10^{-3}	0.752×10^{-1}	0.100×10^{-3}	0.389
15	0.558×10^{-3}	0.790×10^{-3}	0.680×10^{-1}	0.138	0.634
16	0.131×10^{-2}	0.159×10^{-2}	0.226	0.721×10^{-2}	0.274×10^{-1}
17	0.139×10^{-1}	0.112×10^{-1}	0.204×10^{-1}	0.218×10^{-2}	0.570×10^{-2}
18	0.134×10^{-2}	0.167×10^{-2}	0.268×10^{-1}	0.233×10^{-2}	0.830×10^{-2}
19	0.925×10^{-3}	0.139×10^{-2}	0.363×10^{-1}	0.127×10^{-1}	0.898×10^{-1}
20	0.266×10^{-2}	0.451×10^{-2}	0.320×10^{-1}	0.102×10^{-1}	0.406×10^{-1}
21	0.237×10^{-1}	0.330×10^{-1}	0.967×10^{-2}	0.152×10^{-2}	0.400×10^{-2}
22	0.840×10^{-2}	0.119×10^{-1}	0.290×10^{-1}	0.413×10^{-2}	0.128×10^{-1}
23	0.449×10^{-2}	0.846×10^{-2}	0.150×10^{-1}	0.321×10^{-2}	0.100×10^{-1}
24	0.319×10^{-1}	0.127	0.141×10^{-1}	0.295×10^{-2}	0.748×10^{-2}
25	0.557×10^{-2}	0.130×10^{-1}	0.129×10^{-1}	0.342×10^{-2}	0.960×10^{-2}
26	0.239×10^{-2}	0.461×10^{-2}	0.155×10^{-1}	0.566×10^{-2}	0.176×10^{-1}

TABLE 4.14 (Continued)

Region of Origin	Region of Destination 16	17	18	19	20
1	0.809×10^{-2}	0.683×10^{-2}	0.110×10^{-1}	0.140×10^{-2}	0.469×10^{-4}
2	0.172×10^{-1}	0.121×10^{-1}	0.198×10^{-1}	0.194×10^{-2}	0.562×10^{-4}
3	0.843×10^{-2}	0.185×10^{-2}	0.612×10^{-2}	0.778×10^{-3}	0.174×10^{-4}
4	0.562×10^{-1}	0.130×10^{-1}	0.394×10^{-1}	0.291×10^{-2}	0.672×10^{-4}
5	0.142×10^{-1}	0.188×10^{-1}	0.222×10^{-1}	0.238×10^{-2}	0.864×10^{-4}
6	0.149×10^{-1}	0.351×10^{-1}	0.245×10^{-1}	0.180×10^{-2}	0.624×10^{-4}
7	0.188×10^{-1}	0.293×10^{-1}	0.268×10^{-1}	0.120×10^{-2}	0.353×10^{-4}
8	0.878×10^{-2}	0.151×10^{-1}	0.159×10^{-1}	0.171×10^{-2}	0.707×10^{-4}
9	0.921×10^{-2}	0.391×10^{-1}	0.190×10^{-1}	0.137×10^{-2}	0.552×10^{-4}
10	0.392×10^{-2}	0.736×10^{-2}	0.799×10^{-2}	0.886×10^{-3}	0.418×10^{-4}
11	0.317×10^{-2}	0.120×10^{-1}	0.735×10^{-2}	0.593×10^{-3}	0.276×10^{-4}
12	0.620×10^{-2}	0.155×10^{-1}	0.148×10^{-1}	0.144×10^{-2}	0.755×10^{-4}
13	0.381×10^{-1}	0.122×10^{-2}	0.102×10^{-1}	0.162×10^{-2}	0.232×10^{-4}
14	0.144×10^{-1}	0.154×10^{-2}	0.105×10^{-1}	0.671×10^{-2}	0.875×10^{-4}
15	0.195×10^{-1}	0.144×10^{-2}	0.133×10^{-1}	0.169×10^{-1}	0.124×10^{-3}
16	0.189	0.716×10^{-2}	0.207	0.818×10^{-2}	0.121×10^{-3}
17	0.202×10^{-1}	0.324	0.714×10^{-1}	0.202×10^{-2}	0.799×10^{-4}
18	0.916×10^{-1}	0.112×10^{-1}	0.610	0.468×10^{-2}	0.128×10^{-3}
19	0.309×10^{-1}	0.271×10^{-2}	0.400×10^{-1}	0.667	0.116×10^{-2}
20	0.282×10^{-1}	0.660×10^{-2}	0.677×10^{-1}	0.716×10^{-1}	0.446
21	0.845×10^{-2}	0.520×10^{-1}	0.306×10^{-1}	0.165×10^{-2}	0.921×10^{-4}
22	0.377×10^{-1}	0.592×10^{-1}	0.298	0.740×10^{-2}	0.439×10^{-3}
23	0.146×10^{-1}	0.111×10^{-1}	0.584×10^{-1}	0.699×10^{-2}	0.926×10^{-3}
24	0.104×10^{-1}	0.197×10^{-1}	0.298×10^{-1}	0.309×10^{-2}	0.213×10^{-3}
25	0.993×10^{-2}	0.774×10^{-2}	0.282×10^{-1}	0.531×10^{-2}	0.655×10^{-3}
26	0.110×10^{-1}	0.418×10^{-2}	0.249×10^{-1}	0.128×10^{-1}	0.330×10^{-2}

TABLE 4.14 (Continued)

Region of Origin	Region of Destination					
	21	22	23	24	25	26
1	0.167×10^{-1}	0.136×10^{-1}	0.246×10^{-2}	0.191×10^{-2}	0.973×10^{-4}	0.162×10^{-3}
2	0.194×10^{-1}	0.190×10^{-1}	0.281×10^{-2}	0.169×10^{-2}	0.937×10^{-4}	0.165×10^{-3}
3	0.310×10^{-2}	0.426×10^{-2}	0.660×10^{-3}	0.286×10^{-3}	0.207×10^{-4}	0.450×10^{-4}
4	0.158×10^{-1}	0.232×10^{-1}	0.292×10^{-2}	0.121×10^{-2}	0.815×10^{-4}	0.162×10^{-3}
5	0.470×10^{-1}	0.311×10^{-1}	0.530×10^{-2}	0.499×10^{-2}	0.208×10^{-3}	0.301×10^{-3}
6	0.487×10^{-1}	0.322×10^{-1}	0.418×10^{-2}	0.340×10^{-2}	0.137×10^{-3}	0.192×10^{-3}
7	0.186×10^{-1}	0.214×10^{-1}	0.211×10^{-2}	0.106×10^{-2}	0.559×10^{-4}	0.897×10^{-4}
8	0.551×10^{-1}	0.278×10^{-1}	0.511×10^{-2}	0.696×10^{-2}	0.219×10^{-3}	0.269×10^{-3}
9	0.879×10^{-1}	0.349×10^{-1}	0.461×10^{-2}	0.556×10^{-2}	0.158×10^{-3}	0.184×10^{-3}
10	0.427×10^{-1}	0.171×10^{-1}	0.357×10^{-2}	0.784×10^{-2}	0.173×10^{-3}	0.175×10^{-3}
11	0.749×10^{-1}	0.180×10^{-1}	0.281×10^{-2}	0.621×10^{-2}	0.108×10^{-3}	0.102×10^{-3}
12	0.167	0.409×10^{-1}	0.853×10^{-2}	0.398×10^{-1}	0.410×10^{-3}	0.318×10^{-3}
13	0.212×10^{-2}	0.431×10^{-2}	0.653×10^{-3}	0.190×10^{-3}	0.175×10^{-4}	0.460×10^{-4}
14	0.396×10^{-2}	0.730×10^{-2}	0.166×10^{-2}	0.475×10^{-3}	0.551×10^{-4}	0.200×10^{-3}
15	0.370×10^{-2}	0.804×10^{-2}	0.184×10^{-2}	0.427×10^{-3}	0.549×10^{-4}	0.221×10^{-3}
16	0.110×10^{-1}	0.332×10^{-1}	0.377×10^{-2}	0.833×10^{-3}	0.798×10^{-4}	0.194×10^{-3}
17	0.191	0.147	0.806×10^{-2}	0.446×10^{-2}	0.176×10^{-3}	0.208×10^{-3}
18	0.176×10^{-1}	0.116	0.667×10^{-2}	0.106×10^{-2}	0.100×10^{-3}	0.194×10^{-3}
19	0.811×10^{-2}	0.247×10^{-1}	0.683×10^{-2}	0.938×10^{-3}	0.161×10^{-3}	0.854×10^{-3}
20	0.279×10^{-1}	0.902×10^{-1}	0.557×10^{-1}	0.398×10^{-2}	0.123×10^{-2}	0.136×10^{-1}
21	0.458	0.194	0.180×10^{-1}	0.275×10^{-1}	0.431×10^{-3}	0.302×10^{-3}
22	0.286	0.100×10^{-3}	0.105	0.115×10^{-1}	0.900×10^{-3}	0.930×10^{-3}
23	0.904×10^{-1}	0.358	0.312	0.136×10^{-1}	0.534×10^{-2}	0.284×10^{-2}
24	0.446	0.126	0.438×10^{-1}	0.100×10^{-3}	0.269×10^{-2}	0.101×10^{-2}
25	0.696×10^{-1}	0.988×10^{-1}	0.171	0.268×10^{-1}	0.460	0.582×10^{-2}
26	0.223×10^{-1}	0.466×10^{-1}	0.416×10^{-1}	0.459×10^{-2}	0.265×10^{-2}	0.700

TABLE 4.15 - TRANSITION MATRIX OF PROBABILITY OF MIGRATION BETWEEN REGIONS, 1951-1961

Region of Origin	Region of Destination 1	2	3	4	5
1	0.374	0.261×10^{-1}	0.965×10^{-3}	0.198×10^{-1}	0.217
2	0.803×10^{-3}	0.303	0.435×10^{-2}	0.107	0.514×10^{-1}
3	0.117×10^{-3}	0.171×10^{-1}	0.374	0.179	0.728×10^{-2}
4	0.584×10^{-4}	0.102×10^{-1}	0.435×10^{-2}	0.606	0.475×10^{-2}
5	0.340×10^{-2}	0.260×10^{-1}	0.938×10^{-3}	0.252×10^{-1}	0.130
6	0.359×10^{-3}	0.260×10^{-1}	0.814×10^{-3}	0.334×10^{-1}	0.796×10^{-1}
7	0.856×10^{-4}	0.117×10^{-1}	0.100×10^{-2}	0.101	0.980×10^{-2}
8	0.484×10^{-3}	0.722×10^{-2}	0.424×10^{-3}	0.124×10^{-1}	0.170
9	0.629×10^{-4}	0.240×10^{-2}	0.136×10^{-3}	0.522×10^{-2}	0.157×10^{-1}
10	0.415×10^{-4}	0.851×10^{-3}	0.659×10^{-4}	0.199×10^{-2}	0.789×10^{-2}
11	0.412×10^{-4}	0.127×10^{-2}	0.956×10^{-4}	0.339×10^{-2}	0.819×10^{-2}
12	0.688×10^{-4}	0.191×10^{-2}	0.170×10^{-3}	0.559×10^{-2}	0.112×10^{-1}
13	0.187×10^{-4}	0.150×10^{-2}	0.902×10^{-3}	0.273×10^{-1}	0.146×10^{-2}
14	0.409×10^{-4}	0.205×10^{-2}	0.775×10^{-3}	0.143×10^{-1}	0.293×10^{-2}
15	0.334×10^{-4}	0.174×10^{-2}	0.588×10^{-3}	0.137×10^{-1}	0.260×10^{-2}
16	0.159×10^{-4}	0.116×10^{-2}	0.343×10^{-3}	0.173×10^{-1}	0.144×10^{-2}
17	0.269×10^{-4}	0.167×10^{-2}	0.173×10^{-3}	0.917×10^{-2}	0.363×10^{-2}
18	0.240×10^{-4}	0.151×10^{-2}	0.295×10^{-3}	0.144×10^{-1}	0.246×10^{-2}
19	0.168×10^{-4}	0.832×10^{-3}	0.205×10^{-3}	0.617×10^{-2}	0.147×10^{-2}
20	0.293×10^{-4}	0.128×10^{-2}	0.250×10^{-3}	0.773×10^{-2}	0.276×10^{-2}
21	0.440×10^{-4}	0.186×10^{-2}	0.202×10^{-3}	0.803×10^{-2}	0.606×10^{-2}
22	0.219×10^{-4}	0.110×10^{-2}	0.160×10^{-3}	0.674×10^{-2}	0.251×10^{-2}
23	0.146×10^{-4}	0.612×10^{-3}	0.935×10^{-4}	0.326×10^{-2}	0.159×10^{-2}
24	0.523×10^{-4}	0.174×10^{-2}	0.199×10^{-3}	0.665×10^{-2}	0.676×10^{-2}
25	0.314×10^{-4}	0.112×10^{-2}	0.163×10^{-3}	0.512×10^{-2}	0.338×10^{-2}
26	0.692×10^{-4}	0.261×10^{-2}	0.457×10^{-3}	0.132×10^{-1}	0.659×10^{-2}

TABLE 4.15 (Continued)

Region of Origin	Region of Destination 6	7	8	9	10
1	0.449×10^{-1}	0.161×10^{-2}	0.708×10^{-2}	0.612×10^{-1}	0.537×10^{-4}
2	0.101	0.677×10^{-2}	0.326×10^{-2}	0.720×10^{-1}	0.340×10^{-4}
3	0.124×10^{-1}	0.228×10^{-2}	0.753×10^{-3}	0.160×10^{-1}	0.103×10^{-4}
4	0.123×10^{-1}	0.557×10^{-2}	0.533×10^{-3}	0.150×10^{-1}	0.758×10^{-5}
5	0.156	0.288×10^{-2}	0.389×10^{-1}	0.239	0.159×10^{-3}
6	0.100×10^{-3}	0.788×10^{-2}	0.692×10^{-2}	0.468	0.623×10^{-4}
7	0.526×10^{-1}	0.374	0.121×10^{-2}	0.558×10^{-1}	0.171×10^{-4}
8	0.529×10^{-1}	0.156×10^{-2}	0.100×10^{-3}	0.290	0.760×10^{-3}
9	0.603×10^{-1}	0.108×10^{-2}	0.437×10^{-2}	0.637	0.512×10^{-4}
10	0.603×10^{-2}	0.249×10^{-3}	0.859×10^{-2}	0.384×10^{-1}	0.173
11	0.134×10^{-1}	0.549×10^{-3}	0.427×10^{-2}	0.198	0.159×10^{-3}
12	0.138×10^{-1}	0.754×10^{-3}	0.580×10^{-2}	0.995×10^{-1}	0.532×10^{-3}
13	0.266×10^{-2}	0.653×10^{-3}	0.199×10^{-3}	0.465×10^{-2}	0.333×10^{-5}
14	0.397×10^{-2}	0.597×10^{-3}	0.439×10^{-3}	0.804×10^{-2}	0.788×10^{-5}
15	0.382×10^{-2}	0.628×10^{-3}	0.411×10^{-3}	0.811×10^{-2}	0.773×10^{-5}
16	0.303×10^{-2}	0.891×10^{-3}	0.224×10^{-3}	0.609×10^{-2}	0.410×10^{-5}
17	0.128×10^{-1}	0.259×10^{-2}	0.718×10^{-3}	0.436×10^{-1}	0.141×10^{-4}
18	0.544×10^{-2}	0.140×10^{-2}	0.438×10^{-3}	0.133×10^{-1}	0.887×10^{-5}
19	0.231×10^{-2}	0.383×10^{-3}	0.262×10^{-3}	0.559×10^{-2}	0.546×10^{-5}
20	0.417×10^{-2}	0.595×10^{-3}	0.554×10^{-3}	0.115×10^{-1}	0.130×10^{-4}
21	0.126×10^{-1}	0.127×10^{-2}	0.168×10^{-2}	0.662×10^{-1}	0.501×10^{-4}
22	0.522×10^{-2}	0.861×10^{-3}	0.543×10^{-3}	0.173×10^{-1}	0.132×10^{-4}
23	0.259×10^{-2}	0.334×10^{-3}	0.369×10^{-3}	0.875×10^{-2}	0.100×10^{-4}
24	0.968×10^{-2}	0.811×10^{-3}	0.218×10^{-2}	0.464×10^{-1}	0.912×10^{-4}
25	0.469×10^{-2}	0.501×10^{-3}	0.852×10^{-3}	0.165×10^{-1}	0.259×10^{-4}
26	0.891×10^{-2}	0.107×10^{-2}	0.143×10^{-2}	0.264×10^{-1}	0.364×10^{-4}

TABLE 4.15 (Continued)

Region of Origin	Region of Destination 11	12	13	14	15
1	0.128×10^{-1}	0.157×10^{-1}	0.294×10^{-1}	0.405×10^{-2}	0.896×10^{-2}
2	0.121×10^{-1}	0.135×10^{-1}	0.729×10^{-1}	0.628×10^{-2}	0.144×10^{-1}
3	0.360×10^{-2}	0.472×10^{-2}	0.172	0.933×10^{-2}	0.191×10^{-1}
4	0.310×10^{-2}	0.377×10^{-2}	0.127	0.418×10^{-2}	0.108×10^{-1}
5	0.397×10^{-1}	0.401×10^{-1}	0.358×10^{-1}	0.454×10^{-2}	0.109×10^{-1}
6	0.332×10^{-1}	0.253×10^{-1}	0.334×10^{-1}	0.314×10^{-2}	0.817×10^{-2}
7	0.907×10^{-2}	0.919×10^{-2}	0.547×10^{-1}	0.315×10^{-2}	0.896×10^{-2}
8	0.906×10^{-1}	0.906×10^{-1}	0.214×10^{-1}	0.298×10^{-2}	0.753×10^{-2}
9	0.632×10^{-1}	0.234×10^{-1}	0.753×10^{-2}	0.821×10^{-3}	0.224×10^{-2}
10	0.382×10^{-1}	0.940×10^{-1}	0.404×10^{-2}	0.604×10^{-3}	0.160×10^{-2}
11	0.390	0.158	0.625×10^{-2}	0.788×10^{-3}	0.221×10^{-2}
12	0.215	0.260	0.121×10^{-1}	0.175×10^{-2}	0.494×10^{-2}
13	0.123×10^{-2}	0.177×10^{-2}	0.713	0.103×10^{-1}	0.319×10^{-1}
14	0.246×10^{-2}	0.403×10^{-2}	0.163	0.100×10^{-3}	0.594
15	0.255×10^{-2}	0.422×10^{-2}	0.187	0.220	0.280
16	0.171×10^{-2}	0.245×10^{-2}	0.171	0.488×10^{-2}	0.216×10^{-1}
17	0.110×10^{-1}	0.109×10^{-1}	0.155×10^{-1}	0.129×10^{-2}	0.411×10^{-2}
18	0.415×10^{-2}	0.611×10^{-2}	0.605×10^{-1}	0.421×10^{-2}	0.176×10^{-1}
19	0.193×10^{-2}	0.336×10^{-2}	0.512×10^{-1}	0.124×10^{-1}	0.961×10^{-1}
20	0.453×10^{-2}	0.878×10^{-2}	0.418×10^{-1}	0.932×10^{-2}	0.432×10^{-1}
21	0.418×10^{-1}	0.677×10^{-1}	0.186×10^{-1}	0.221×10^{-2}	0.704×10^{-2}
22	0.697×10^{-2}	0.115×10^{-1}	0.210×10^{-1}	0.228×10^{-2}	0.843×10^{-2}
23	0.409×10^{-2}	0.871×10^{-2}	0.120×10^{-1}	0.188×10^{-2}	0.696×10^{-2}
24	0.373×10^{-1}	0.155	0.178×10^{-1}	0.274×10^{-2}	0.845×10^{-2}
25	0.858×10^{-2}	0.222×10^{-1}	0.182×10^{-1}	0.344×10^{-2}	0.116×10^{-1}
26	0.114×10^{-1}	0.247×10^{-1}	0.603×10^{-1}	0.152×10^{-1}	0.565×10^{-1}

TABLE 4.15 (Continued)

Region of Origin	Region of Destination 16	17	18	19	20
1	0.856×10^{-2}	0.119×10^{-1}	0.163×10^{-1}	0.191×10^{-2}	0.166×10^{-3}
2	0.193×10^{-1}	0.228×10^{-1}	0.315×10^{-1}	0.292×10^{-2}	0.223×10^{-3}
3	0.223×10^{-1}	0.931×10^{-2}	0.242×10^{-1}	0.283×10^{-2}	0.172×10^{-3}
4	0.274×10^{-1}	0.120×10^{-1}	0.287×10^{-1}	0.207×10^{-2}	0.129×10^{-3}
5	0.121×10^{-1}	0.251×10^{-1}	0.260×10^{-1}	0.261×10^{-2}	0.245×10^{-3}
6	0.130×10^{-1}	0.454×10^{-1}	0.294×10^{-1}	0.210×10^{-2}	0.189×10^{-3}
7	0.254×10^{-1}	0.611×10^{-1}	0.505×10^{-1}	0.232×10^{-2}	0.179×10^{-3}
8	0.823×10^{-2}	0.217×10^{-1}	0.202×10^{-1}	0.203×10^{-2}	0.214×10^{-3}
9	0.337×10^{-2}	0.199×10^{-1}	0.929×10^{-2}	0.654×10^{-3}	0.673×10^{-4}
10	0.170×10^{-2}	0.483×10^{-2}	0.464×10^{-2}	0.480×10^{-3}	0.568×10^{-4}
11	0.295×10^{-2}	0.157×10^{-1}	0.906×10^{-2}	0.706×10^{-3}	0.827×10^{-4}
12	0.576×10^{-2}	0.212×10^{-1}	0.181×10^{-1}	0.167×10^{-2}	0.217×10^{-3}
13	0.585×10^{-1}	0.435×10^{-2}	0.260×10^{-1}	0.370×10^{-2}	0.151×10^{-3}
14	0.264×10^{-1}	0.578×10^{-2}	0.288×10^{-1}	0.142×10^{-1}	0.532×10^{-3}
15	0.433×10^{-1}	0.679×10^{-2}	0.446×10^{-1}	0.408×10^{-1}	0.912×10^{-3}
16	0.455	0.893×10^{-2}	0.162	0.664×10^{-2}	0.278×10^{-3}
17	0.108×10^{-1}	0.530	0.484×10^{-1}	0.147×10^{-2}	0.149×10^{-3}
18	0.128	0.317×10^{-1}	0.341	0.958×10^{-2}	0.699×10^{-3}
19	0.314×10^{-1}	0.574×10^{-2}	0.571×10^{-1}	0.254	0.324×10^{-2}
20	0.264×10^{-1}	0.116×10^{-1}	0.835×10^{-1}	0.651×10^{-1}	0.374
21	0.116×10^{-1}	0.961×10^{-1}	0.532×10^{-1}	0.289×10^{-2}	0.398×10^{-3}
22	0.188×10^{-1}	0.456×10^{-1}	0.173	0.468×10^{-2}	0.680×10^{-3}
23	0.827×10^{-2}	0.104×10^{-1}	0.413×10^{-1}	0.458×10^{-2}	0.137×10^{-2}
24	0.961×10^{-2}	0.276×10^{-1}	0.357×10^{-1}	0.348×10^{-2}	0.577×10^{-3}
25	0.101×10^{-1}	0.131×10^{-1}	0.373×10^{-1}	0.619×10^{-2}	0.173×10^{-2}
26	0.314×10^{-1}	0.214×10^{-1}	0.942×10^{-1}	0.385×10^{-1}	0.208×10^{-1}

TABLE 4.15 (Continued)

Region of Origin	Region of Destination 21	22	23	24	25	26
1	0.218×10^{-1}	0.179×10^{-1}	0.355×10^{-2}	0.244×10^{-2}	0.159×10^{-3}	0.263×10^{-3}
2	0.285×10^{-1}	0.277×10^{-1}	0.458×10^{-2}	0.251×10^{-2}	0.175×10^{-3}	0.306×10^{-3}
3	0.122×10^{-1}	0.160×10^{-1}	0.275×10^{-2}	0.112×10^{-2}	0.999×10^{-4}	0.210×10^{-3}
4	0.117×10^{-1}	0.163×10^{-1}	0.233×10^{-2}	0.914×10^{-3}	0.764×10^{-4}	0.148×10^{-3}
5	0.470×10^{-1}	0.321×10^{-1}	0.602×10^{-2}	0.492×10^{-2}	0.267×10^{-3}	0.391×10^{-3}
6	0.501×10^{-1}	0.342×10^{-1}	0.502×10^{-2}	0.360×10^{-2}	0.189×10^{-3}	0.270×10^{-3}
7	0.336×10^{-1}	0.376×10^{-1}	0.432×10^{-2}	0.201×10^{-2}	0.135×10^{-3}	0.216×10^{-3}
8	0.569×10^{-1}	0.304×10^{-1}	0.612×10^{-2}	0.696×10^{-2}	0.294×10^{-3}	0.371×10^{-3}
9	0.338×10^{-1}	0.146×10^{-1}	0.218×10^{-2}	0.223×10^{-2}	0.858×10^{-4}	0.103×10^{-3}
10	0.192×10^{-1}	0.835×10^{-2}	0.188×10^{-2}	0.328×10^{-2}	0.101×10^{-3}	0.107×10^{-3}
11	0.668×10^{-1}	0.184×10^{-1}	0.320×10^{-2}	0.561×10^{-1}	0.140×10^{-3}	0.139×10^{-3}
12	0.147	0.411×10^{-1}	0.924×10^{-2}	0.316×10^{-3}	0.491×10^{-3}	0.410×10^{-3}
13	0.587×10^{-2}	0.109×10^{-1}	0.185×10^{-2}	0.529×10^{-2}	0.584×10^{-4}	0.145×10^{-3}
14	0.110×10^{-1}	0.189×10^{-1}	0.459×10^{-2}	0.129×10^{-2}	0.175×10^{-3}	0.582×10^{-3}
15	0.130×10^{-1}	0.258×10^{-1}	0.631×10^{-2}	0.147×10^{-3}	0.219×10^{-3}	0.800×10^{-3}
16	0.108×10^{-1}	0.287×10^{-1}	0.374×10^{-2}	0.835×10^{-2}	0.954×10^{-4}	0.222×10^{-3}
17	0.108	0.843×10^{-1}	0.573×10^{-2}	0.291×10^{-2}	0.150×10^{-3}	0.183×10^{-3}
18	0.390×10^{-1}	0.209	0.148×10^{-1}	0.246×10^{-2}	0.278×10^{-3}	0.528×10^{-3}
19	0.126×10^{-1}	0.338×10^{-1}	0.978×10^{-2}	0.143×10^{-2}	0.275×10^{-3}	0.129×10^{-2}
20	0.349×10^{-1}	0.984×10^{-1}	0.585×10^{-1}	0.475×10^{-2}	0.155×10^{-2}	0.140×10^{-1}
21	0.191	0.254	0.276×10^{-1}	0.349×10^{-1}	0.788×10^{-3}	0.601×10^{-3}
22	0.154	0.364	0.566×10^{-1}	0.674×10^{-2}	0.644×10^{-3}	0.696×10^{-3}
23	0.564×10^{-1}	0.191	0.189	0.806×10^{-2}	0.326×10^{-2}	0.195×10^{-2}
24	0.371	0.119	0.420×10^{-1}	0.100×10^{-3}	0.278×10^{-2}	0.121×10^{-2}
25	0.773×10^{-1}	0.104	0.156	0.256×10^{-1}	0.374	0.640×10^{-2}
26	0.786×10^{-1}	0.150	0.124	0.149×10^{-1}	0.853×10^{-2}	0.973×10^{-1}

TABLE 4.16 - POPULATION DECREASE RATES USED IN THE MARKOV CHIAN MODEL++

Region	D Rates (Percentage)			
	1921-1931+	1931-1941	1941-1951	1951-1961+
1	0.00	0.00	14.28*	9.05
2	9.38	5.68	8.16*	9.05
3	0.00	4.00	66.67	9.05
4	9.38	6.76	6.85	9.05
5	9.38	6.38	7.91	9.05
6	9.38	4.57	8.32	9.05
7	0.00	20.83	42.10*	9.05
8	9.38	7.09	13.72*	9.05
9	9.38	7.40	8.34	9.05
10	0.00	65.10*	0.00	90.00*
11	9.38	5.97	7.85	9.05
12	9.38	10.07	9.70	9.05
13	9.38	6.31	8.15	9.05
14	9.38	10.21*	47.53*	9.05
15	9.38	8.25	8.44	9.05
16	9.38	5.05	24.42*	9.05
17	9.38	5.92	7.52	9.05
18	9.38	8.54	8.53	9.05
19	9.38	3.97	6.60	40.70*
20	0.00	0.00	7.89	9.05
21	9.38	7.76	8.33	9.05
22	9.38	7.29	8.30	9.05
23	9.38	5.90	6.48	43.30*
24	9.38	4.35	7.38	9.05
25	0.00	4.17	5.00	9.05
26	0.00	1.99	6.72	9.05

*The values are equivalent to the percentage population decrease over the period. The corresponding X values have been taken as zero.

+Birth and Death rates are not known by regions in these decades. The average rate for the study area has been applied to the population of individual regions.

++The percentage figure is obtained by expressing the absolute decrease as a percentage of the population at the beginning of the time period.

TABLE 4.17 - POPULATION INCREASE VALUES USED IN THE MARKOV CHAIN MODEL

Region	1921-1931	1931-1941	1941-1951	1951-1961
1	24	4	0*	88
2	421	172	337	1160
3	24	83	0	12
4	325	169	136	391
5	1619	485	1291	519
6	547	110	413	190
7	24	115	0*	45
8	89	26	0*	183
9	2476	961	1419	1010
10	10	0*	13	0*
11	1100	190	498	403
12	277	35	527	1598
13	3324	1250	1251	5706
14	1525	0*	0*	120
15	2102	327	536	1200
16	1086	780	0*	166
17	332	416	1385	356
18	1333	362	583	1045
19	1116	435	88	0*
20	33	9	92	196
21	980	357	531	910
22	743	408	435	1096
23	1722	715	430	0*
24	1071	319	568	34
25	24	69	22	2
26	193	250	25	191

*A zero value indicates that the region has a net population decrease. The decrease has been added into the D value of the transition matrices.

TABLE 4.18 - MARKOV CHAIN EQUILIBRIUM SOLUTIONS AND ACTUAL DISTRIBUTIONS: PERCENTAGE POPULATION DISTRIBUTIONS BY TIME PERIOD*

Region	Percentage Population Distribution							
	1921-1931		1931-1941		1941-1951		1951-1961	
	Equilibrium	Actual	Equilibrium	Actual	Equilibrium	Actual	Equilibrium	Actual
1	0.022	0.065	0.026	0.066	0.027	0.058	0.097	0.216
2	0.373	1.281	0.465	1.989	1.021	2.510	1.377	4.256
3	0.030	0.067	0.204	0.219	0.103	0.113	0.120	0.103
4	0.975	1.117	2.777	1.716	2.097	1.755	4.616	1.812
5	1.845	7.615	1.001	7.426	2.172	8.377	1.201	7.824
6	0.961	1.357	2.555	1.413	2.403	1.681	1.718	1.590
7	0.022	0.065	0.237	0.268	0.284	0.186	0.265	0.214
8	0.187	0.381	0.195	0.479	0.328	0.417	0.300	0.650
9	12.445	8.926	7.477	9.107	7.772	10.219	11.865	8.143
10	0.009	0.027	0.006	0.000	0.023	0.021	0.006	0.000
11	1.899	2.983	2.870	2.936	5.236	3.428	4.073	3.365
12	2.439	3.617	1.892	3.416	3.203	3.428	3.853	6.254
13	17.252	19.241	17.462	17.085	22.945	16.975	26.081	22.866
14	2.042	5.465	2.106	4.965	1.111	3.329	1.479	2.229
15	13.212	9.649	6.684	8.804	5.395	7.496	4.323	6.401
16	3.404	3.736	6.471	4.215	3.789	3.253	6.501	2.298
17	3.455	1.915	4.267	2.468	4.617	3.319	5.004	2.980
18	5.701	6.509	9.849	6.229	13.100	6.669	8.269	6.050
19	0.856	4.553	1.781	6.807	1.382	4.373	0.746	2.992
20	0.030	0.086	0.057	0.089	0.173	0.204	0.245	0.535
21	9.198	5.737	9.259	5.590	11.828	5.349	6.165	5.573
22	19.224	6.808	11.852	6.483	7.001	6.106	9.487	5.777
23	3.571	4.114	9.014	4.864	2.474	4.744	1.418	3.836
24	0.683	4.213	0.847	4.335	1.224	3.833	0.573	2.650
25	0.022	0.065	0.140	0.188	0.103	0.200	0.051	0.142
26	0.140	0.407	0.404	0.839	0.187	0.856	0.168	1.242

*The values refer to the percentage of the total population of the area in each region.

prevent the occurrence of negative values in the X vector (Table 4.16). The X vectors are normally calculated by adding each region's expected external migration gain or loss to the natural increase.

The inverse of the transition matrix (N, the fundamental matrix) multiplied by the X vector gives the expected equilibrium distribution of population (Table 4.17). These equilibrium distributions are tested against real observations for each time period in Chapter Five.

CHAPTER FIVE

TESTING OF THE MODELS

A theory is necessarily based on more or less true approximating assumptions. The derivation of assumptions is technically referred to as the problem of specification (Bartlett 1962, p. 24). In this chapter the specification or statistical model is tested against the statistical data. The theoretical distribution is related to the observed facts to assist in the interpretation of the facts and to facilitate further inference.

What are the theoretical distributions? In the Markov chain model, they are the equilibrium solutions derived for each time period. In the Monte Carlo model, the simulation results are assumed to be representative of an underlying theoretical distribution. These theoretical distributions have to be related to actual observations. If it is assumed that the real world has only a probability existence, then the observed facts can be considered as one possible sample drawn from many. Consequently, an adequate one-sample test is desired, which gauges the goodness-of-fit of the sample to the theoretical population (Siegel 1956, p. 35). Such tests indicate whether a particular sample (here, real-world observations) could have come from some specified theoretical population (the simulation and equilibrium solutions).

A common technique to test the difference between the observed (sample) statistics and the expected (population) parameters involves the use of a "t" test. Such a test requires the assumption that the sample

observations come from a normally distributed population. (Siegel 1956, p. 35). Since this assumption of normality can not be made here, the one-sample test used in this study must be nonparametric rather than one of the generally more powerful parametric varieties.

Chi-Square and Kolmogorov-Smirnov tests are frequently used as one-sample tests. Siegel (1956, p. 59) considered that the Kolmogorov-Smirnov test is a more powerful goodness-of-fit test than Chi-Square. The latter can not cope with small class values, so that class intervals have to be altered, resulting in loss of information. Small values, or even zero entries, do not affect the performance of the Kolmogorov-Smirnov test. Consequently, the latter technique has been used in this study to estimate the goodness-of-fit between real-world and theoretical distributions.

The Kolmogorov-Smirnov test involves the specification of theoretical and sample cumulative frequences. The degree of agreement between the expected and observed values is determined by the maximum deviation between the two,

$$D = \text{maximum } \left| F_o(X) - S_n(X) \right| \qquad (5.1)$$

where D is the maximum difference, $F_o(X)$ the theoretical cumulative distribution, and $S_n(X)$ the observed cumulative distribution. The null hypothesis, H_o, is that difference between the two distributions are not significant but due merely to chance. The smaller the value of D, the more likely that the null hypothesis can not be rejected. Since the sampling distribution of D under H_o is known (Siegel 1956, Table E, p. 251), comparison of the D value with critical values in the sampling

distribution permits acceptance or rejection of the null hypothesis. A full description of the Kolmogorov-Smirnov technique is provided by Siegel (1958, p. 47-52). The test should only be used when the variable to be examined has a continuous distribution. However, migration data are discrete. If the test is applied to discrete data, then the resultant error is on the "safe" side, since the test is a conservative one (Siegel 1956, p. 59).

In this study, the real-world observations are tested against theoretical and simulated distributions. Single simulation results are considered as samples drawn from averaged distributions to determine if one simulation run varies greatly from the averaged result. To facilitate testing of observed and expected distributions by the Kolmogorov-Smirnov technique, the data are presented in three ways:

1. the absolute population values for each region are measured in nine size groups;

2. percentage population values for each region are distributed among eight categories;

3. the net population gain or loss of a region over a time period is expressed as a percentage of the population of the region at the beginning of a time period. The resultant values are arranged in nine classes.

In all three cases, the number of classes used for measurement are as numerous as possible, in order to take full advantage of the sensitivity of the Kolmogorov-Smirnov technique. (Tables 5.1, 5.2, 5.3).

To illustrate the technique, Table 5.4 shows the expected (equilibrium) and observed (real world) cumulative frequency distributions

TABLE 5.1 - FREQUENCY DISTRIBUTIONS OF STUDY REGIONS WITH POPULATION BY SIZE CLASSES

Class	1921-1931				1931-1941			
	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation
< 50	6	6	6	6	3	3	3	3
50- 100	0	2	0	0	2	3	2	2
100- 250	2	1	2	2	2	3	2	2
250- 500	2	4	1	1	1	2	2	2
500-1000	2	4	3	3	3	3	2	2
1000-2000	6	3	6	6	5	4	5	5
2000-4000	7	2	7	7	9	5	8	8
4000-8000	1	4	1	1	1	3	1	1
8000>	0	0	0	0	0	0	1	1

TABLE 5.1 (Continued)

Class	1941-1951				1951-1961			
	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation
< 50	2	2	3	3	1	2	1	1
50- 100	4	4	2	3	2	3	3	2
100- 250	1	2	2	1	2	3	1	3
250- 500	1	1	1	1	2	2	2	1
500-1000	2	3	2	2	2	5	2	3
1000-2000	6	6	5	5	7	0	8	7
2000-4000	7	5	8	8	7	7	5	5
4000-8000	2	2	2	2	2	3	3	3
8000>	1	1	1	1	1	1	1	1

TABLE 5.2 - FREQUENCY DISTRIBUTIONS OF STUDY REGIONS

BY PERCENTAGE OF TOTAL POPULATION

Class*	1921-1931				1931-1941			
	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation
< 0.25	6	8	6	6	5	7	5	5
0.25- 0.50	2	1	2	2	2	2	2	2
0.50- 1.00	0	4	0	0	1	1	1	1
1.00- 2.00	4	2	4	4	3	3	3	3
2.00- 4.00	3	5	5	5	3	4	3	3
4.00- 8.00	8	1	6	6	9	4	9	9
8.00-16.00	2	3	2	2	2	4	2	2
16.00>	1	2	1	1	1	1	1	1

*The values refer to the percentage of the total population in the area.

TABLE 5.2 (Continued)

Class*	1941-1951				1951-1961			
	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation
< 0.25	6	6	6	6	5	6	5	5
0.25- 0.50	1	2	1	1	0	2	1	1
0.50- 1.00	1	0	1	1	2	2	2	2
1.00- 2.00	2	4	2	2	3	5	2	2
2.00- 4.00	6	6	5	5	7	1	8	8
4.00- 8.00	7	5	7	7	7	6	6	6
8.00-16.00	2	2	3	3	1	3	1	1
16.00>	1	1	1	1	1	1	1	1

*The values refer to the percentage of the total population in the area.

TABLE 5.3 - FREQUENCY DISTRIBUTIONS OF STUDY REGIONS GAIN OR LOSS BY PERCENTAGE POPULATION

Class*	1921-1931				1931-1941			
	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation
< -50	0	0	0	0	1	0	1	1
-15 - -50	0	3	0	0	0	5	1	1
0 - -15	0	3	0	0	0	1	4	4
0 - 10	7	7	7	7	6	0	3	3
10 - 25	0	0	0	0	9	4	6	7
25 - 50	0	0	1	1	4	4	5	5
50 - 100	6	2	5	5	2	2	2	1
100 - 200	4	3	3	3	1	4	1	1
200 >	9	8	10	10	3	6	3	3

*The values refer to percentage net gain or loss of population over the period, where the absolute net gain or loss is expressed as a percentage of the population at the beginning of the time period.

TABLE 5.3 (Continued)

Class*	1941-1951				1951-1961			
	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation	Real World	Equilibrium Solution	Averaged Simulation	Single Simulation
< -50	0	0	2	2	1	1	1	1
-15 - -50	3	6	4	4	3	9	4	4
0 - -15	3	4	1	2	4	2	2	2
0 - 10	6	1	6	6	2	2	1	2
10 - 25	6	3	4	3	6	2	5	4
25 - 50	5	3	5	4	3	5	4	4
50 - 100	2	4	2	3	4	3	2	3
100 - 200	1	3	2	2	1	1	6	5
200 >	0	2	0	0	2	1	1	1

*The values refer to percentage net gain or loss of population over the period, where the absolute net gain or loss is expressed as a percentage of the population at the beginning of the time period.

for regions falling within certain percentage size classes. The maximum D value in this example is 5/26 (0.1923). To cause rejection of

TABLE 5.4 - EXPECTED AND OBSERVED CUMULATIVE FREQUENCY DISTRIBUTIONS: EQUILIBRIUM AND REAL WORLD PERCENTAGE DISTRIBUTIONS, 1921-1931

Class (Percentage)	Equilibrium Solution	Observed Real World	$F_o(X)$	$S_{26}(X)$	$D=\lvert F_o(X)-S_{26}(X)\rvert$
- 0.25	8	6	8/26	6/26	2/26
0.25 - 0.50	1	2	9/26	8/26	1/26
0.50 - 1.00	4	0	13/26	8/26	5/26
1.00 - 2.00	2	4	15/26	12/26	3/26
2.00 - 4.00	5	3	20/26	15/26	5/26
4.00 - 8.00	1	8	21/26	23/26	2/26
8.00 - 16.00	3	2	24/26	25/26	1/26
16.00 -	2	1	26/26	26/26	0

the null hypothesis—there is no significant difference between the equilibrium (expected) and real world (observed) distributions other than chance variations—the D value must exceed the critical probability value chosen.

The maximum D value is less than the p value associated with the 0.20 level, (Table 5.5). In other words, the probability of any difference between expected and observed distributions being due to chance factors alone is greater than 20 chances in 100. The null hypothesis can not be rejected. The observed distribution of regions can be considered as a sample of the expected equilibrium solution.

TABLE 5.5 - CRITICAL VALUES OF D IN THE KOLMOGOROV-SMIRNOV ONE-SAMPLE TEST

Degrees of Freedom	Level of Significance D = maximum $\lvert F_o(x)-S_n(x)\rvert$				
	.20	.15	.10	.05	.01
26	0.2102	0.2240	0.2397	0.2672	0.3202

Source: Adapted from Siegel (1956, Table E, p. 251).

Equilibrium Solutions and Real World Observations:

Comparison of two distributions is very difficult when there are as many as twenty-six observations (Table 4.18). Only the more obvious deviations can be picked out by a visual examination. For example, the Markov chain model predicts that region 22, not region 13, will contain the largest percentage of the total population of the study area in time period one (1921-1931). After the first time period, region 13 is clearly the dominant region in the equilibrium soluitons, as it is in the real world observations. The Kolmogorov-Smirnov test compares the two cumulative frequency distributions on the basis of the maximum observed difference. While some information is lost, because of the necessity to group the regions into size classes, the test takes into account overall variations in the system of regions.

Results of the tests (Table 5.6) indicate that there is a promising relationship between the theoretical equilibrium distributions and the real world observations. The null hypothesis states that there is no significant difference between the expected and observed

TABLE 5.6 - MAXIMUM DIFFERENCE VALUES BETWEEN EQUILIBRIUM AND REAL WORLD CUMULATIVE FREQUENCY DISTRIBUTIONS

Time Period	Frequency Distribution: No. of Regions by Population	No. of Regions by Percentage of Total Population	No. of Regions by Percentage Gain or Loss
1921-1931	0.192	0.192	0.231
1931-1941	0.115	0.115	0.231
1941-1951	0.077	0.077	0.231
1951-1961	0.231	0.192	0.231

distributions, and that any differences are due to chance factors. In no instance can the null hypothesis be rejected at the conventional critical limits of 0.05 and 0.01. Five out of the twelve D values indicate that the null hypothesis may be rejected at the comparatively weak confidence level of 0.15. A value equivalent to or less than the p value of 0.10 suggests that any differences, occurring purely by chance and not significant may arise 10 times in 100 cases.

TABLE 5.7 - MAXIMUM DIFFERENCE VALUES BETWEEN SIMULATED AND REAL WORLD CUMULATIVE FREQUENCY DISTRIBUTIONS

Time Period	Frequency Distribution: No. of Regions by Population		No. of Regions by Percentage of Total Population		No. of Regions by Percentage Gain or Loss	
	Averaged	Single	Averaged	Single	Averaged	Single
1921-1931	0.038	0.038	0.077	0.077	0.038	0.038
1931-1941	0.038	0.038	0	0	0.192	0.192
1941-1951	0.038	0.038	0.038	0.038	0.115	0.115
1951-1961	0.038	0.038	0.038	0.038	0.154	0.115

Simulated Results and Real World Observations

Pitts (1963) suggested that the results of a series of simulations can be averaged to produce a combined or "omega" distribution. Since individual simulation runs vary widely, the averaged solution is assumed to have more theoretical significance than one single simulation. In this study, observations are tested against a randomly selected simulation in each time period, and against the average of fifty simulations. Kolmogorov-Smirnov tests indicate a high degree of association between real-world observations and both averaged and single simulation runs, (Table 5.7). In no instance can the null hypothesis be rejected with any significant degree of confidence. All maximum difference values are less than the critical 0.20 confidence level.

TABLE 5.8 - MAXIMUM DIFFERENCE VALUES BETWEEN AVERAGED AND SINGLE SIMULATIONS

	Frequency Distribution		
Time Period	No. of Regions by Population	No. of Regions by Percentage of Total Population	No. of Regions by Percentage Gain or Loss
1921-1931	0	0	0
1931-1941	0	0	0.038
1941-1951	0.038	0	0.038
1951-1961	0.038	0	0.038

There is very little difference between the results of the single simulations and the average of fifty runs (Tables 4.10 and 4.11). When both distributions were grouped into classes, the frequencies were

very close, and identical in some instances (Tables 5.1, 5.2, and 5.3). Kolmogorov-Smirnov tests were applied to these frequencies to see if one single simulation varied significantly from the combination of runs. The single simulations (observations) were regarded as samples coming from the averaged (theoretical) solution. The resultant maximum difference values are extremely low (Table 5.8). Had different single simulations been chosen, the results might have been quite different, as evidenced by the standard deviations of the combined simulation runs (Table 4.9). However, the single simulations were chosen completely at random, and must be considered typical.

TABLE 5.9 - MAXIMUM DIFFERENCE VALUES BETWEEN EQUILIBRIUM AND SIMULATED DISTRIBUTIONS

	Frequency Distribution					
Time Period	No. of Regions by Population		No. of Regions by Percentage of Total Population		No. of Regions by Percentage Gain or Loss	
	Averaged	Single	Averaged	Single	Averaged	Single
1921-1931	0.192	0.192	0.192	0.192	0.231	0.231
1931-1941	0.115	0.115	0.115	0.115	0.231	0.269
1941-1951	0.115	0.115	0.115	0.115	0.192	0.154
1951-1961	0.231	0.192	0.192	0.192	0.231	0.192

Equilibrium Solutions and Simulated Results

The experimental sampling model (Monte Carlo) and the analytic Markov chain model were tested against each other to see if further differences could be revealed. In general, results suggest little

difference between the equilibrium (expected) and simulated (observed) distributions (Table 5.9). However, when the equilibrium solution is compared with the results of a single simulation in time period two (1931-1941), in one instance, the null hypothesis of only chance differences between expected and observed distributions can be rejected at the 0.05 confidence level. The corresponding equilibrium and averaged solutions have a much smaller maximum difference value, so that the hypothesis can only be rejected at the 0.15 confidence level. If the 0.05 confidence level is regarded as critical, then combined and single simulation results perform quite differently. This occurred only once in twelve possibilities at the 0.05 confidence level, but suggests that greater reliability can be given to combined rather than single simulations.

Summary

Regional population figures have been organized into frequency distributions. Similar frequencies have been derived for the regional scores predicted by the combined and single simulations, and the Markovian equilibrium solutions. The expected (theoretical) and observed (real-world) frequencies are tested by a one-sample Kolmogorov-Smirnov test. The one-sample test permits the acceptance or rejection of the hypothesis, that there is no significant difference between observed and expected frequencies, other than chance variations. The real-world observations can never be rejected, at the 0.05 confidence level as coming, from the predicted, theoretical distributions. The Kolmogorov-Smirnov technique also indicates little difference between the results of single and combined simulations.

Analysis and evaluation of the Monte Carlo and Markovian models, based on the results of the Kolmogorov-Smirnov test, follow in Chapter Six.

CHAPTER SIX

CONCLUSIONS

The primary aim of the thesis is to investigate the relative efficiency of the Monte Carlo simulations and Markovian equilibrium solutions. Out of this has emerged the secondary task of evaluating the significance of an averaged simulation solution and a single simulation run. On the basis of the Kolmogorov-Smirnov tests presented in Chapter Five, the Monte Carlo and Markovian models successfully reproduce the features of the real world within reasonable bounds of confidence. Qualifications must be added, since the one-sample Kolmogorov-Smirnov test is less powerful than its parametric counterparts. Also, as regions' scores were grouped into classes, a certain amount of information was lost. Despite these limitations, the results clearly indicate that both the Monte Carlo and the Markovian techniques function reasonably well.

The maximum difference values between simulated (expected) and real world (observed) cumulative frequencies (Table 5.7) are generally less than the difference values obtained by comparison of the equilibrium solutions with the real world (Table 5.6). However, in the latter case, the null hypothesis—differences between observed and expected distributions are not significant, but due to chance—can never be rejected at the critical 0.05 and 0.01 confidence levels.

The Kolmogorov-Smirnov test examines the total variation of the system of regions. Since the test requires the specification of cumulative frequencies, positive and negative scores of regions can not be

evaluated by this method. The tabulated data (Tables 4.10, 4.11, 4.18) and choropleth maps (Figures 14-21) indicate some of the more significant deviations of observed values from expected figures. Table 4.18 presents the variation between equilibrium (expected) and real world (observed) values for regions, in terms of the percentage of the total population of the area in each region. In general, there is a reasonable fit, though the Markovian model in time period one (1921-1931) predicts that region 22 will be dominant, not region 13, which contains Grande Prairie, the capital seat of the Peace District. The Markovian model continually under-predicts the percentage value of region 5. Deviations between the predicted simulated values and the real world observations are also apparent (Figures 14-21). For example, both the single run and combined simulations exaggerate the size of region 13 in time period two, 1931-1941 (Tables 4.10, 4.11). Such fluctuations are generally less than those observed between equilibrium and real world values.

A cursory interpretation of the results would conclude that the simulation model performs slightly better than the Markovian model, since the overall variation between observed and expected values is less, and the prediction of the scores of individual regions is closer to the observed figures. However, certain qualifications must be made.

One reason for smaller deviations of simulated values from observations may lie in the method by which the constants of natural increase and external migration are introduced into the Monte Carlo model. Such constants impose rigid controls upon the results of the simulations. In the Markovian model, following derivation of X and D vectors, which take natural increase and gain or loss due to external migration into

account, the whole population was subjected to the process of population redistribution. Calculation of the expected number of potential migrants was not necessary for the Markovian model, but was required by the Monte Carlo method. This is a powerful factor in favour of the Markov chain model of migration, since the elements of population change—natural increase, and external and internal migration—occur together continuously in the model.

Hence, the Markovian model is more realistic than the Monte Carlo method. Moreover, the operational procedure of the former is simpler. The number of calculations required by the Monte Carlo model is much greater than the Markovian techniques. The running and averaging in each time period of fifty simulations required many hours of computing time. In contrast, once the transition matrices and the X vectors were constructed, the Markovian equilibrium solutions were achieved by a simple procedure of matrix and vector multiplication.

The author contends that the operation of the Markovian model of migration is more satisfying than the Monte Carlo technique. Admittedly, the Kolmogorov-Smirnov tests indicate a slightly closer fit between observed and expected values when the Monte Carlo model is used. Yet, the results achieved by the Markovian model are still excellent, and in no instance can the null hypothesis, that both observed and expected distributions are similiar apart from random deviations, be rejected at the 0.05 confidence level. The implementation of the Markov chain model is both more realistic, since the elements of population change occur together continuously, and more flexible, as the computations

required are much less than those demanded by the Monte Carlo model.

The excellent performance of the Markov chain migration model should not be allowed to obscure several drawbacks, especially if the technique is tested with more detailed information. Both the Markovian and Monte Carlo models in this study are general in nature. In both models a simple inverse relationship between migration intensity and distance is the basis on which a theoretical population distribution is generated. The under-prediction by the Markovian model of the size of region 13 in time period one, and the continual under-estimation of the size of region 5, indicate that factors other than the simple inverse-distance relationship need to be considered. The low predicted value for region 13 in time period one would suggest that the marginal location of this region is enough to off-set the high population density (Figure 9). In contrast, region 2 received an exaggerated value, because of its central location, and its moderately high population density. Again, the constant under-prediction of region 5 by the Markovian model seems to be a reflection of that region's location in the south-east extremity of the study area (Figure 9). Since the transition matrices, derived by means of a gravity model, incorporate both distances between, and size of, regions relative location becomes extremely important. Some additional functional weighting is desired in the case of region 13 because Grande Prairie is a regional centre for the Peace District. However, the equilibrium solution only under-estimated the size of region 13 in time period one (1921-1931). In the other three decades, the size in population of region 13 was sufficient to overcome the disadvantages of the region's location.

If independent variables other than distance were considered necessary in a model of migration, then how are they to be incorporated? In Geography, analytic models have not been too successful because the derivation of joint probability functions for spatial variables poses many mathematical problems (King 1969, p. 230). The use of physical and functional barriers in simulation methods is a response to the absence of joint probability functions. In the Monte Carlo model barriers were used to modify the inverse-distance postulate. No similar procedure was possible in the Markov chain model. Additional factors could perhaps be incorporated in matrix operation as recommended by Rogers (1966b). He included age and sex structure in a model of migration. In theory, two-dimensional matrices could be expanded to three or even q-dimensions to accommodate many more than two variables (Olsson and Gale 1968). However, whether complex multi-dimensional Markovian models can be made operational remains to be seen.

The Monte Carlo method has the powerful advantage of being able to incorporate chance deviations from the expected probabilities of migration. This is extremely important, owing to the disadvantages imposed by the probabilities. The gravity model used to estimate the Markovian transition matrices, and the Pareto functions employed by the Monte Carlo method, have similar disadvantages. Both incorporate only one independent variable (distance) and both require estimation of parameters from the observations to be predicted. Such parameters have little theoretical significance. These drawbacks can be overcome somewhat by the Monte Carlo method since the process of allocating random numbers to migrations permits chance factors to modify the expected results. The

incorporation of the chance element overcomes, to a certain degree, the ignorance of the researcher because it reflects unforeseen factors. In the Markovian model, the probabilities of migration were basically similar to those utilized in the Monte Carlo model, but were not modified to include chance variations. Consequently, several significant deviations between predicted and observed values arose, such as the under-estimation of the size of region 5.

As a general model of migration, the Markovian technique is successful, but if the method were applied to situations where more detailed data are available, and if a greater level of precision were required, then the transition probabilities would need to include the effect of other significant physical, social, and economic variables.

Markovian models are proposed as an analytic alternative to simulation methods, but Monte Carlo models can also be improved. The Monte Carlo model produced similar results for both single and combined simulations. As a further check, simulation distributions were compared to the equilibrium solutions to reveal any unsuspected differences (Table 5.9). In one case, the results of a single simulation were rejected at the 0.05 confidence level as coming from the theoretical population. The averaged solution was not rejected. The author believes that an averaged result is more accurate than a single simulation in the case where discrete moving units are considered. High speed computers make the averaging of many simulation runs simple, although the task is still time consuming.

Since the Markov chain model of migration used in this study

reasonably duplicates the performance of the Monte Carlo model, and conceptually is much neater and not so laborious to operate, Markovian models could be given wider application. In Geography, Markovian models have chiefly treated population movements (Rogers 1966a, Compton 1969) and diffusion (Brown 1963). Economists have used the technique to predict the interregional movement of industry (Hampton 1968) and trade (Smith 1961). Conceivably, Markovian models could investigate a wide variety of forms of spatial interaction, including migration, diffusion, spread of settlement, consumer behaviour, and changes in urban morphology.

The technique need not be applied solely to the prediction of future spatial distributions. The models in this study were applied to a wide historical time span. The continuity of results through time suggests that the Markovian model could be used as an historical - predictive device in a detailed study. Like the Monte Carlo method, the model could test hypotheses about spatial behaviour through time.

If Markovian models are to be given wider application, then the transition matrices must be meaningful, and capable of incorporating the influence of several significant variables. Unless these conditions can be fulfilled, the conceptual elegance of Markovian models will be void and Monte Carlo techniques will continue to function as a reasonable alternative to oversimplfied analytic models.

BIBLIOGRAPHY

Ajo, R. 1953. Contributions to Social Physics. Lund Studies in Geography: Series B, Human Geography 11.

Alberta, Department of Public Health and Vital Statistics. Annual Reports, 1921 to 1961, Edmonton.

Alcaly, R.E. 1967. Aggregation and Gravity Models; Some Empirical Evidence. Journal of Regional Science 7: 61-73.

Anderson, I.B. 1966. Internal Migration in Canada, 1921-1961. Economic Council of Canada, Staff Study No. 13, Ottawa: Roger Duhamel.

Bartlett, M.S. 1962. Essays on Probability and Statistics. London: Methuen.

Bergsten, K.E. 1951. Variability in Intensity of Urban Fields as Illustrated by Birth-Places. Lund Studies in Geography: Series B, Human Geography 3: 25-32.

Berry, B.J.L., and Schwind, P.J. 1969. Information and Entropy in Migrant Flows. Geographical Analysis 1: 5-14.

Bharucha-Reid, A.T. 1960. Elements of the Theory of Markov Processes and their Applications. New York: McGraw-Hill.

Bowes, G.E., ed. 1963. Peace River Chronicles. Vancouver: Prescott Publishing Company.

Brown, L.A. 1963. The Diffusion of Innovation; A Markov Chain Approach. Department of Geography, Discussion Paper No. 3, Evanston: Northwestern University.

__________ 1968a. Diffusion Dynamics; A Review and Revision of the Quantitative Theory of the Spatial Diffusion of Innovation. Lund Studies in Geography: Series B, Human Geography 29.

__________ 1968b. Diffusion Processes and Location; A Conceptual Framework and Bibliography. Bibliography Series No. 4, Philadelphia: Regional Science Research Institute.

Bunge, W. 1969. Simplicity. Geographical Analysis 1: 388-391.

Bylund, E. 1960. Theoretical Considerations Regarding the Distribution of Settlement in Inner North Sweden. Geografiska Annaler 42: 255-231.

Canada, Dominion Bureau of Statistics. Census of Canada, 1921, 1931, 1941, 1951, 1956, 1961, Ottawa.

Canada, Dominion Bureau of Statistics. Census of Prairie Provinces, 1916, 1926, 1936, 1946, Ottawa.

Carder, A.C. 1965. Climate in the Upper Peace River Region, Canada. Publication No. 1224, Ottawa: Canada, Department of Agriculture.

Casetti, E. 1969. Why Do Diffusion Processes Conform to Logistic Trends? Geographical Analysis 1: 101-105.

Chorley, R.J., and Haggett, P., eds. 1967. Models in Geography. London: Methuen.

Churchman, C.W., Ackoff, R.L., and Arnoff, E.L. 1957. Introduction to Operations Research. New York: John Wiley and Sons.

Clark, W.A.V. 1965. Markov-Chain Analysis; An Application to the Movement of Rental Housing Areas. Annals of the Association of American Geographers 55: 351-359.

Cole, J.P., and King, C.A.M. 1968. Quantitative Geography; Techniques and Theories in Geography. London: John Wiley and Sons.

Colenutt, R.J. 1969. Linear Diffusion in an Urban Setting. Geographical Analysis 1: 106-114.

Compton, P.A. 1969. Internal Migration and Population Change in Hungary Between 1959 and 1965. Institute of British Geographers, Publications 47: 111-130.

Curry, L. 1966. Chance and Landscape. In Northern Geographical Essays in Honour of G.H.J. Daysh, ed. J.W. House, pp. 40-55. Newcastle-upon-Tyne: Oriel Press.

________ 1967. Central Places in the Random Spatial Economy. Journal of Regional Science 7: 217-238.

Czamanski, S. 1964. A Model of Urban Growth. Regional Science Association, Papers and Proceedings 13: 170-220.

Dacey, M.F. 1966. A Compound Probability Law for a Pattern More Dispersed than Random and with Areal Inhomogeneity. Economic Geography 42: 172-179.

De Cani, J.S. 1961. On the Construction of Stochastic Models of Population Growth and Migration. Journal of Regional Science 3: 1-13.

Dodd, S.C. 1955. Diffusion is Predictable; Testing Probability Models for Laws of Interaction. American Sociological Review 20: 392-401.

Feller, W. 1957. An Introduction to Probability Theory and its Applications. New York: John Wiley and Sons.

Garrison, W.L. 1962. Towards Simulation Models of Urban Growth and Development. Lund Studies in Geography: Series B, Human Geography 24: 92-108.

Goldberg, M.A. 1968. Bay Area Simulation Study; Employment Location Models. Annals of Regional Science 2: 161-173.

Golledge, R.G. 1967. A Conceptual Framework of a Market Decision Process. Journal of Regional Science 7: 239-258.

Golledge, R.G., and Brown, L.A. 1967. Search, Learning, and the Market Decision Process. Geografiska Annaler: Series B, Human Geography 49: 116-124.

Gould, D. 1969. Spatial Diffusion. Commission on College Geography, Research Paper No. 4, Washington: Association of American Geographers.

Hägerstrand, T. 1951. Migration and the Growth of Cultural Regions. Lund Studies in Geography: Series B, Human Geography 3: 33-36.

______________ 1952. Propagation of Innovation Waves. Lund Studies in Geography: Series B, Human Geography 4: 3-19.

______________ 1957. Migration and Area; Survey of a Sample of Swedish Migration Fields and Hypothetical Considerations on Their Genesis. Lund Studies in Geography: Series B, Human Geography 13: 27-158.

______________ 1965. A Monte Carlo Approach to Diffusion. Achives Europeenes de Sociologie 6: 43-67.

______________ 1966. Aspects of the Spatial Structure of Social Communication and the Diffusion of Information. Regional Science Association, Papers and Proceedings 16: 27-42.

______________ 1967a. Innovation Diffusion as a Spatial Process. Chicago: University of Chicago Press.

______________ 1967b. On Monte Carlo Simulation of Diffusion. In Quantitative Geography, Part I; Economic and Cultural Topics, eds. W.L. Garrison, and D.F. Marble, pp. 1-32. Evanston: Northwestern University.

Hammersley, J.M., and Handscomb, D.C. 1964. Monte Carlo Methods. London: Methuen.

Hampton, P. 1968. Regional Economic Development in New Zealand. Journal of Regional Science 8: 41-55.

Harling, J. 1958. Simulation Techniques in Operations Research; A Review. Operations Research 6: 307-319.

Harris, C.C. 1968. A Stochastic Process Model of Residentical Development. Journal of Regional Science 8: 29-39.

Harvey, D. 1967. Models of the Evaluation of Spatial Patterns in Human Geography. In Models in Geography, eds. R.J. Chorley, and P. Haggett, pp. 549-668. London: Methuen.

Hudson, J.C. 1969a. A Location Theory for Rural Settlement. Annals of the Association of American Geographers 59: 365-381.

__________ 1969b. Diffusion in a Central Place System. Geographical Analysis 1: 45-58.

Isard, W. 1960. Methods of Regional Analysis; An Introduction to Regional Science. New York: John Wiley and Sons.

Kemeny, J.G. 1959. Finite Mathematical Structures. Englewood Cliffs, New Jersey: Prentice-Hall.

Kemeny, J.G., and Snell, J.L. 1960. Finite Markov Chains. Princeton: Van Nostrand.

__________ 1962. Mathematical Models in the Social Sciences. Boston: Ginn.

Kemeny, J.G., Snell, J.L., and Knapp, A.W. 1966. Denumerable Markov Chains. Princeton: Van Nostrand.

Kemeny, J.G., Snell, J.L., and Thompson, G.L. 1956. Introduction to Finite Mathematics. Englewood Cliffs, New Jersey: Prentice-Hall.

King, G. 1953. The Monte Carlo Method as a Natural Mode of Expression in Operations Research. Operations Research 1: 46-51.

King, G.W. 1951. Stochastic Method in Statistical Mechanics. In Monte Carlo Method, United States, National Bureau of Standards, Applied Mathematics, Series 12, pp. 12-18. Washington.

King, L.J. 1969. Statistical Analysis in Geography. Englewood Cliffs, New Jersey: Prentice-Hall.

Kitto, F.H. 1930. The Peace River Country, Canada. Ottawa: Canada, Department of the Interior.

Kulldorff, G. 1955. Migration Probabilities. Lund Studies in Geography; Series B, Human Geography 14.

Lövgren, E. 1957. Mutual Relations Between Migration Fields. Lund Studies in Geography: Series B, Human Geography 13: 159-169.

Lycan, R. 1969. Interprovincial Migration in Canada; The Role of Spatial and Economic Factors. Canadian Geographer 13: 237-254.

Maass, A., et al. 1962. Design of Water Resource Systems. Cambridge, Massachusetts: Harvard University Press.

MacGregor, J.G. 1952. Land of Twelve Foot Davis. Edmonton: Applied Art Products.

MacKintosh, W.A. 1934. Prairie Settlement, The Geographical Setting. In Canadian Frontiers of Settlement Series, eds. W.A. MacKintosh and W.L.G. Joerg. Toronto: The Macmillian Company of Canada.

Malm, R., Olsson, G., and Wärneryd, D. 1966. Approaches to Simulation of Urban Growth. Geografiska Annaler: Series B, Human Geography 48: 9-22.

Marble, D.F. 1964. Two Computer Programs for the Analysis of Simple Markov Chains. Discussion Paper No. 6, Philadelphia: Regional Science Research Institute.

Marble, D.F., and Bowlby, S.R. 1968. Computer Programs for the Operational Analysis of Hägerstrand Type Spatial Diffusion Models. Office of Naval Research, Geography Branch, Spatial Diffusion Study, Technical Report No. 9, Evanston: Department of Geography, Northwestern University.

Marble, D.F., and Nystuen, J.D. 1963. An Approach to the Direct Measurement of Community Mean Information Fields. Regional Science Association, Papers and Proceedings 11: 99-109.

Marriott, P.J. 1969. Migration of People To and Within the County of Grande Prairie, Alberta, 1956 to 1967. Unpublished M.A. thesis: University of Alberta.

Meyer, H.A., ed. 1956. Symposium on Monte Carlo Methods. New York: John Wiley and Sons.

Morrill, R.L. 1960. Simulation of Central Place Patterns Over Time. Lund Studies in Geography: Series B, Human Geography 24: 109-120.

____________ 1963a. The Distribution of Migration Distances. Regional Science Association, Papers and Proceedings 11: 75-84.

____________ 1963b. The Development of Spatial Distributions of Towns in Sweden; An Historical—Predictive Approach. Annals of the Association of American Geographers 53: 1-14.

____________ 1965a. Expansion of the Urban Fringe; A Simulation Experiment. Regional Science Association, Papers and Proceedings 15: 185-202.

____________ 1965b. Migration and the Spread and Growth of Urban Settlement. Lund Studies in Geography: Series B, Human Geography 26.

____________ 1965c. The Negro Ghetto; Problems and Alternatives. Geographical Review 55: 339-361.

____________ 1968. Waves of Spatial Diffusion. Journal of Regional Science 8: 1-18.

Morrill, R.L., and Pitts, F.R. 1967. Marriage, Migration, and the Mean Information Field; A Study in Uniqueness and Generality. Annals of the Association of American Geographers 57: 401-422.

Myles, E.L. 1965. The Emperor of Peace River. Edmonton: Institute of Applied Art.

Nelson, P. 1959. Migration, Real Income, and Information. Journal of Regional Science 1: 43-74.

Nystuen, J.D. 1967. A Theory and Simulation of Intraurban Travel. In Quantitative Geography, Part I; Economic and Cultural Topics, eds. W.L. Garrison, and D.F. Marble, pp. 54-84. Evanston: Northwestern University.

Olsson, G. 1965a. Distance and Human Interaction; A Migration Study. Geografiska Annaler: Series B, Human Geography 47: 3-43.

__________ 1965b. Distance and Human Interaction; A Review and Bibliography. Bibliography Series No. 2, Philadelphia: Regional Science Research Institute.

Olsson, G., and Gale, S. 1968. Spatial Theory and Human Behaviour. Regional Science Association, Papers and Proceedings 21: 229-242.

Pitts, F.R. 1962. Chorology Revisited - Computerwise. Professional Geographer 14: 8-12.

__________ 1963. Problems in Computer Simulation of Diffusion. Regional Science Association, Papers and Proceedings 11: 111-119.

__________ 1965. HAGER III and HAGER IV; Two Monte Carlo Computer Programs for the Study of Spatial Diffusion Problems. Office of Naval Research, Geography Branch, Spatial Diffusion Study, Technical Report No. 4, Evanston: Department of Geography, Northwestern University.

__________ 1967. MIFCAL and NONCEL; Two Computer Programs for the Generation of Hägerstrand Models to an Irregular Lattice. Office of Naval Research, Geography Branch, Spatial Diffusion Study, Technical Report No. 7, Evanston: Department of Geography, Northwestern University.

Porter, R. 1956. Approach to Migration Through its Mechanism. Geografiska Annaler 38: 317-343.

Pred, A. 1967. Postscript. In Innovation Diffusion as a Spatial Process, T. Hägerstrand, pp. 299-324. Chicago: University of Chicago Press.

Raroport, A. 1966. What is Information? In Communication and Culture; Readings in the Codes of Human Interaction, ed. A.G. Smith, pp. 41-55. New York: Holt, Rinehart, and Winston.

Recht, J.R. 1968. Bay Area Simulation Study; Residential Location Model. Annals of Regional Science 2: 142-152.

Reilly, W.J. 1931. The Law of Retail Gravitation. New York: The Knickerbocker Press.

Rogers, A. 1966a. A Markovian Policy Model of Interregional Migration. Regional Science Association, Papers and Proceedings 17: 205-224.

__________ 1966b. Matrix Analysis of Interregional Population Growth and Distribution. Regional Science Association, Papers and Proceedings 18: 177-196.

Rogers, E.M. 1962. Diffusion of Innovations. New York: The Free Press of Glencoe.

Rushton, G. 1969. Analysis of Spatial Behaviour by Revealed Space Preferences. Annals of the Association of American Geographers 59: 391-400.

Schreider, Yu. A. 1966. The Monte Carlo Method. Oxford: Pergamon Press.

Shubik, M., ed. 1964. Game Theory and Related Approaches to Social Behaviour. New York: John Wiley and Sons.

Siegel, S. 1956. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill.

Smith, A.G. ed. 1966. Communication and Culture; Readings in the Codes of Human Interaction. New York: Holt, Rinehart, and Winston.

Smith, P.E. 1961. Markov Chains, Exchange Matrices, and Regional Development. Journal of Regional Science 3: 27-36.

Spiegel, M.R. 1961. Theory and Problems of Statistics. New York: Schaum Publishing Company.

Stouffer, S.A. 1960. Intervening Opportunities and Competing Migrants. Journal of Regional Science 2: 1-26.

Tobler, W.R. 1967. Of Maps and Matrices. Journal of Regional Science 7: 275-280.

Tocher, K.D. 1963. The Art of Simulation. London: The English Universities Press.

Tolley, G.S. 1963. Population Adjustment and Economic Activity; Three Studies. Regional Science Association, Papers and Proceedings 11: 85-97.

Tracie, C.J. 1967. Agricultural Settlement in the South Peace River Area. Unpublished M.A. thesis: University of Alberta.

__________ 1970. An Analysis of Three Variables Affecting Farm Location in the Process of Agricultural Settlement; The South Peace River Area. Unpublished Ph.D. thesis: University of Alberta.

United States, Department of Commerce. 1951. Monte Carlo Method. National Bureau of Standards, Applied Mathematics, Series No. 12: Washington.

Vanderhill, B.G. 1963. Trends in the Peace River Country. Canadian Geographer 7: 33-41.

Wallace, J.N. 1929. The Wintering Partners on Peace River. Ottawa: Thorburn and Abbott.

Walter, G.R. 1968. Bay Area Simulation Study; Economic and Population Models. Annals of Regional Science 2: 59-71.

Weaver, W. 1966. The Mathematics of Communication. In Communication and Culture; Readings in the Codes of Human Interaction, ed. A.G. Smith, pp. 15-24. New York: Holt, Rinehart, and Winston.

Wellman, B.B. 1965. A History of the Peace River Country of Western Canada. Unpublished M.A. thesis: University of Arizona.

Wendel, B. 1957. Regional Aspects of Internal Migration and Mobility in Sweden, 1946-1950. Lund Studies in Geography: Series B, Human Geography 13: 7-26.

Willis, G.A. 1966. Development of Transportation in the Peace River Region of Alberta and British Columbia, with an Evaluation of Present Day Rail and Road Commodity Flow Patterns. Unpublished M.A. thesis: University of Alberta.

Wolpert, J. 1965. Behavioural Aspects of the Decision to Migrate. Regional Science Association, Papers and Proceedings 15: 159-172.

Yuill, R.S. 1964. A Simulation Study of Barrier Effects in Spatial Diffusion Problems. Office of Naval Research, Geography Branch, Spatial Diffusion Study, Technical Report No. 1, Evanston: Department of Geography, Northwestern University.

Zelinsky, W. 1962. A Bibliographic Guide to Population Geography. Department of Geography, Research Paper No. 80, Chicago: University of Chicago.

Zipf, G.K. 1949. Human Behavior and the Principle of Least Effort. Cambridge, Massachusetts: Addison-Wesley.

APPENDIX A

CALCULATION OF PARETO FUNCTIONS

A number of studies has shown the existence of a significant inverse relationship between migration intensity and distance (Hägerstrand 1957, Morrill 1963a, and Olsson 1965a). This relationship was identified for migration in the Peace District 1921 to 1961, and expressed in mathematical form in the Monte Carlo model.

The first step was the definition of the two variables, migration intensity and distance. Imagine a chart with a point of origin, around which a series of rings are drawn at ten kilometres apart. The distance variables were taken as the distance (kilometres) between the point of origin and the mid-point of each ring. The number of migrations ending in each ring was obtained to provide the migration intensity variables, expressed as the number of migrations per square kilometre to each ring.

The variables were tabulated for each time period (Tables A-1 to A-4) and a general decrease of migration with distance was observed. Following identification of the inverse-distance relationship, the association of the two variables was expressed mathematically by fitting Pareto functions to the data.

$$M = \frac{c}{D^b}$$

TABLE A.1 - OBSERVED LOCAL MIGRATION IN THE PEACE DISTRICT, 1921-1931*

Ring Area (Square Kilometres)	Number of Migrants	Number of Migrants per Sq. Kilometre	Distance from Origin to Ring Centre (Kilometres)
314	16	0.05095	5
942	4	0.00425	15
1884	6	0.00318	25
3140	1	0.00032	35
4710	4	0.00085	45
6594	2	0.00030	55
8792	0	0.00000	65
11304	1	0.00009	75

TABLE A.2 - OBSERVED LOCAL MIGRATION IN THE PEACE DISTRICT, 1931-1941*

Ring Area (Square Kilometres)	Number of Migrants	Number of Migrants per Sq. Kilometre	Distance from Origin to Ring Centre (Kilometres)
314	34	0.10828	5
942	11	0.01168	15
1884	7	0.00371	25
3140	2	0.00064	35
4710	6	0.00127	45
6594	2	0.00030	55
8792	1	0.00011	65
11304	5	0.00044	75
14129	1	0.00007	85
17270	1	0.00006	95

*Source: Sample survey by the author 1969.

TABLE A.3 - OBSERVED LOCAL MIGRATION IN THE PEACE DISTRICT, 1941-1951*

Ring Area (Square Kilometres)	Number of Migrants	Number of Migrants per Square Kilometre	Distance from Origin to Ring Centre (Kilometres)
314	24	0.07643	5
942	19	0.02017	15
1884	27	0.01423	25
3140	6	0.00191	35
4710	0	0.00000	45
6594	6	0.00091	55
8792	4	0.00045	65
11304	4	0.00035	75
14130	2	0.00014	85
17270	0	0.00000	95
20724	2	0.00009	105

TABLE A.4 - OBSERVED LOCAL MIGRATION IN THE PEACE DISTRICT, 1951-1961*

Ring Area (Square Kilometres)	Number of Migrants	Number of Migrants per Square Kilometre	Distance from Origin to Ring Centre (Kilometres)
314	25	0.07962	5
942	13	0.01380	15
1884	12	0.00637	25
3140	5	0.00159	35
4710	6	0.00127	45
6594	4	0.00061	55
8792	3	0.00034	65
11304	9	0.00080	75
14130	5	0.00035	85

*Source: Sample survey by the author, 1969.

or

$$M = cD^{-b}$$

where M is the number of migrants per square kilometre, D is the distance variable, and c and b are constants to be estimated. The function should be fitted to data which exhibit a linear relationship. To make this assumption more reasonable, the variables were transformed to $\log M$ and $\log D$. The function was fitted by the method of least squares. The least squares line has the form,

$$Y = a_o + a_1X$$

where Y is the dependent variable, X is the independent variable, a_o is the slope, and a_1 is the intercept. Knowing,

$$M = cD^{-b}$$

and

$$\log M = \log c - b \log D$$

assume that $Y = \log M$ and $X = \log D$, then

$$Y = a_o + a_1X$$

where $a_o = \log c$ and $a_1 = -b$. These constants can be estimated by the procedure suggested by Spiegel (1961, p. 220),

$$a_o = \frac{(\sum y)(\sum x^2) - (\sum y)(\sum xy)}{N\sum x^2 - (\sum x)^2}$$

and

$$a_1 = \frac{N\sum xy - (\sum x)(\sum y)}{N\sum x^2 - (\sum x)^2}$$

where N is the number of observations. By this method, the following equations were determined,

1921 - 1931	$M = 5.370D^{-2.80}$
1931 - 1941	$M = 1.047D^{-2.46}$
1941 - 1951	$M = 4.787D^{-2.77}$
1951 - 1961	$M = 0.6918D^{-2.18}$

The values of c and b were fed directly into the Monte Carlo programme to determine the raw probability of interaction between region i and region j . This is given by

$$I_{ij} = \frac{c(A)}{D^b}$$

where c and b are the derived constants, A is the area of the

destination region, and D is the distance between the centres of the two regions. For example, the raw probability of contact in time period one (1921-1931) between region one and region two (see Table 4.6) is

$$\frac{5.370\ (3339)}{61.474^{2.80}} = 0.1800$$

where 5.370 and 2.80 are the c and b values respectively of the Pareto expression, 3339 is the area (sq. kilometres) of region two, and 61.474 is the distance (kilometres) between the centres of regions one and two.

Raw probabilities were then smoothed and adjusted by population to derive the final probability of migration between region i and region j.

The Pareto function extends from zero to infinity. Consequently, in this thesis, the distances put into the function are greater than zero, and less than infinity (Morrill 1963a, p. 82). Thus, if

$$M = p(d)$$

then

$$p(d) = cD^{-b} \ ; \quad 0 < D < \infty \ .$$

APPENDIX B

MARKOV CHAINS

A simple numerical example of a regular Markov chain is provided to illustrate the comments made in Chapter Three. Assume an initial probability vector $p^{(o)} = (\frac{3}{4}, \frac{1}{4})$, where the probability of phenomena being distributed in state one is 3/4 , and in state two is 1/4 . The transition probabilities are given by

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}$$

How will the phenomena be distributed after one step, knowing the initial distribution and the transition probabilities? The probability of phenomena being in state one after one step is the sum of the probabilities of being in each of the possible states at the previous step, and then moving to state one on this step. Thus,

$$p_1^{(1)} = p_1^{(o)} \cdot p_{11} + p_2^{(o)} \cdot p_{21}$$

where $p_1^{(1)}$ is the probability of phenomena being distributed in state one after one step, $p_1^{(o)}$ and $p_2^{(o)}$ are the probabilities of phenomena being initially distributed in state one and state two, respectively, p_{11} is the transition probability of moving from state one to state one, and p_{21} the probability of moving from state two to state one. Equally,

$$p_2^{(1)} = p_1^{(o)} \cdot p_{12} + p_2^{(o)} \cdot p_{22}$$

where $p_2^{(1)}$ is the probability of phenomena being in state two after one step, and p_{12} and p_{22} are the transition probabilities of moving from state one to state two, and from state two to state two, respectively. In this example,

$$p_1^{(1)} = \frac{3}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{3}{4}$$

$$= \frac{9}{16}$$

and

$$p_1^{(2)} = \frac{3}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{4}$$

$$= \frac{7}{16}$$

In practice, it can be shown that

$$p^{(n)} = p^{(n-1)} \cdot P$$

where $p^{(n)}$ represents the probable distribution of phenomena after n steps, $p^{(n-1)}$ the probable distribution after $n - 1$ steps, and P is the matrix of transition probabilities. In this example,

$$p^{(1)} = \left(\frac{3}{4} \quad \frac{1}{4}\right) \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{2} \end{pmatrix}$$

$$= \left(\frac{9}{16} \quad \frac{7}{16}\right)$$

where the probability of phenomena being in state one after one time period is 9/16 and in state two is 7/16 . Equally,

$$p^{(2)} = p^{(1)} \cdot P$$

$$= \left(\frac{9}{16} \quad \frac{7}{16}\right) \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}$$

$$= \left(\frac{39}{64} \quad \frac{25}{64}\right)$$

where the probability of being in states one and two after two steps is 39/64 and 25/64 , respectively. In all cases, p is a probability vector which sums to unity. Also, if

$$p^{(n)} = p^{(n-1)} \cdot P$$

then

$$p^{(n)} = p^{(o)} \cdot p^{n}$$

where

$$P^n = P \cdot P \cdot P \ldots P \quad (\text{n factors})$$

Thus,

$$p^{(2)} = p^{(o)} \cdot P^2$$

where

$$P^2 = P \cdot P$$

In the above example,

$$p^{(2)} = (\tfrac{3}{4} \cdot \tfrac{1}{4}) \cdot \begin{pmatrix} \frac{5}{8} & \frac{3}{8} \\ \frac{9}{16} & \frac{7}{16} \end{pmatrix}$$

$$= (\tfrac{39}{64} \quad \tfrac{25}{64})$$

This is the same answer derived by expression

$$p^{(n)} = p^{(n-1)} \cdot P$$

Matrix multiplication enables successive state of the system to be computed rapidly.

APPENDIX C

TABLES OF POPULATION STATISTICS BY REGIONS, 1921 to 1961

TABLE C.1 - REGIONAL POPULATION FIGURES, 1921-1931

Region	Population 1921	Births 1921-1931	Deaths 1921-1931	Natural Increase 1921-1931	Expected Total Migration Gain or Loss 1921-1931*	Expected External Migration Gain or Loss 1921-1931*
1	0	0	0	0	24	24
2	148	48	13	35	292	373
3	2	0	0	0	23	24
4	104	34	9	25	285	291
5	1322	433	124	309	1192	1186
6	77	25	7	18	408	522
7	0	0	0	0	24	24
8	72	23	6	17	52	66
9	1156	379	108	271	1882	2097
10	0	0	0	0	10	10
11	251	82	23	59	796	1018
12	1233	404	115	289	-181	-127
13	4030	1321	378	943	2160	2003
14	852	279	79	200	974	1246
15	1271	416	119	297	2009	1686
16	354	116	33	83	948	970
17	430	140	40	100	180	192
18	1260	413	118	295	858	920
19	290	95	27	68	1330	1021
20	0	0	0	0	32	33
21	1337	438	125	313	477	542
22	1456	477	136	341	727	266
23	46	15	4	11	1468	1707
24	359	117	33	84	1119	954
25	0	0	0	0	24	24
26	0	0	0	0	151	193

Source: Census of Canada and Vital Statistics reports of Alberta

*Calculated from data derived by a sample.

TABLE C.2 - REGIONAL POPULATION FIGURES, 1931-1941

Region	Population 1931	Births 1931-1941	Deaths 1931-1941	Natural Increase 1931-1941	Expected Total Migration Gain or Loss 1931-1941*	Expected External Migration Gain or Loss 1931-1941*
1	24	1	0	1	3	3
2	475	172	27	145	226	0
3	25	2	1	1	67	81
4	414	169	28	141	175	0
5	2823	707	180	527	-191	-222
6	503	101	23	78	20	9
7	24	4	5	-1	91	111
8	141	26	10	16	47	0
9	3309	1092	245	847	-282	-131
10	10	0	0	0	-10	-12
11	1106	298	66	232	-89	-108
12	1341	424	135	289	-177	-389
13	7133	1674	450	1224	-1089	-424
14	2026	410	120	290	-204	-497
15	3577	1065	295	770	-602	-738
16	1385	321	70	251	157	459
17	710	252	42	210	130	164
18	2413	856	206	650	-413	-494
19	1688	289	67	222	136	146
20	32	0	0	0	6	9
21	2127	506	165	341	-90	-149
22	2524	674	184	490	-256	-266
23	1525	528	90	438	106	187
24	1562	366	68	298	-16	-47
25	24	5	1	4	52	64
26	151	25	3	22	184	225

Source: Census of Canada and Vital Statistics Reports of Alberta

*Calculated from data derived by a sample.

TABLE C.3 - REGIONAL POPULATION FIGURES, 1941-1951

Region	Population 1941	Births 1941-1951	Deaths 1941-1951	Natural Increase 1941-1951	Expected Total Migration Gain or Loss 1941-1951*	Expected External Migration Gain or Loss 1941-1951*
1	28	8	3	5	-5	-9
2	846	337	69	268	103	0
3	93	14	4	10	-48	-72
4	730	239	50	189	-68	-103
5	3159	1040	250	790	112	251
6	610	215	50	165	49	198
7	114	29	10	19	-43	-67
8	204	58	14	44	-46	-72
9	3874	1356	323	1033	47	63
10	0	0	0	0	10	13
11	1249	399	98	301	112	99
12	1453	527	141	386	356	0
13	7268	2468	592	1876	-915	-1217
14	2112	605	144	461	-959	-1465
15	3745	1218	316	902	-1013	-682
16	1793	552	128	424	-640	-862
17	1050	361	79	282	277	1024
18	2650	942	226	716	-133	-359
19	2046	582	135	447	-373	-494
20	38	7	3	4	57	85
21	2378	769	198	571	-356	-238
22	2758	879	229	650	-448	-444
23	2069	632	134	498	-267	-202
24	1844	568	136	432	-418	0
25	80	18	4	14	3	4
26	357	101	24	77	-19	-76

Source: Census of Canada and Vital Statistics Reports of Alberta

*Calculated from data derived by a sample.

TABLE C.4 - REGIONAL POPULATION FIGURES, 1951-1961

Region	Population 1951	Births 1951-1961	Deaths 1951-1961	Natural Increase 1951-1961	Expected Total Migration Gain or Loss 1951-1961*	Expected External Migration Gain or Loss 1951-1961*
1	28	12	2	10	90	76
2	1217	560	110	450	855	600
3	55	25	4	21	-15	-13
4	851	391	77	314	-91	0
5	4061	1870	367	1503	-928	-1351
6	815	375	73	302	-175	-185
7	90	41	8	33	4	4
8	202	93	18	75	108	90
9	4954	2281	448	1833	-1962	-1271
10	10	4	0	4	-14	-13
11	1662	765	150	615	-283	-362
12	2195	1011	198	813	698	587
13	8229	3790	744	3046	2274	1916
14	1614	743	146	597	-890	-623
15	3634	1673	328	1345	-1186	-473
16	1577	726	142	584	-799	-560
17	1609	741	145	596	-439	-385
18	3233	1489	292	1197	-845	-444
19	2120	976	191	785	-1132	-1648
20	99	45	8	37	181	151
21	2593	1194	234	960	-251	-284
22	2960	1363	267	1096	-633	-267
23	2300	1059	208	851	-878	-1847
24	1858	855	168	687	-975	-821
25	97	44	8	36	-49	-42
26	415	191	37	154	167	0

Source: Census of Canada and Vital Statistics Reports of Alberta

*Calculated from data derived by a sample.

APPENDIX D

THE SAMPLE QUESTIONNAIRE

A 98% response to a questionnaire on population migration was achieved. Interviews were carried out over the telephone. The high degree of success by the telephone method negated the necessity of sending out further questionnaires to be completed by mail. The author believes that a questionnaire administered by telephone is extremely efficient if the circumstances are correct. Firstly, all respondents were sent a covering letter and a list of the questions to be asked. The questions could be asked in a simple direct fashion, and topics of income and age were omitted as they were not considered necessary in this study. The people of this largely rural based area are generally well aware of its recent history, and many had been among the original settlers. Most were more than eager to co-operate and speak of their parent's or their own role in settling this region. In fact, the greatest difficulty encountered was the necessity to break off telephone calls, which were becoming too lengthy. That this technique would work in urban areas is doubtful, since urban residents are usually more aware of the canvassing telephone calls made by salesmen. If the topic had been less interesting to the people of the Peace District, and had the questions been more personal, then the response might not have been so high. The success of the method suggests that geographers might make more use of the telephone to carry out short, non-controversial interviews.

Department of Geography

The University of Alberta
Edmonton, Canada

August, 1969

Dear Sir or Madam:

My name is Ian Lindsay. I am writing a Master's Thesis at the University of Alberta in the Geography Department, and am studying the movement of settlers in the Peace River District.

I have selected your name at random and would like to ask you several questions about the places where you and your parents have lived.

Unfortunately, I will not be able to meet everyone in person, so I will phone you in the next few weeks to obtain your answers to the enclosed questions. I hope this will not inconvenience you.

This telephone conversation will take only a few minutes and will deal only with where you and your family has lived. All information provided will be completely confidential.

Without your assistance, I can not carry out my study of the history of settlement in the Peace River District. I would be grateful for any additional comments you may wish to make.

Sincerely,

Ian Lindsay
Graduate Teaching Assistant
Department of Geography
University of Alberta
Edmonton, Alberta
Phone: 432-4158

IL/amcc
encl.

PEACE RIVER SETTLEMENT QUESTIONNAIRE NO. 1

Number: Respondent:

These questions largely refer to places and dates. Could you please be as precise as possible, especially when referring to the Peace River District. In the latter case, give township, range, and quarter location for rural addresses.

Section I: To be filled in by head of household

1. Where and when was your father born?

   Place:

   Date:

2. Where did your father live before his marriage? Give dates for each address.

3. When were your parents married?

   Date:

4. Where did your father live after his marriage? Include his present address if still alive, and his addresses even if his spouse has died. Give dates for each address.

5. If your father is dead when did he die?

   Date:

6. Where and when was your mother born?

   Place:

   Date:

7. Where did she live before she was married? Give dates for each address.

8. Where did she live after her marriage? Include her present address if still alive, and her addresses even if her spouse has died. Give dates for each address.

9. If your mother is dead, when did she die?

   Date:

10. Where and when were you born?

    Place:

    Date:

11. Where did you live before your marriage? If not married, where have you lived? Give dates for each address.

12. When were you married (if applicable)?

    Date:

13. Where have you lived since your marriage (if applicable)? Give dates for each address.

14. If your spouse is dead when did he or she die?

Date:

Section II: To be filled in by the spouse (if any) of the head of the household

The format of Section II is the same as Section I.

B29957